NEUROETHICS

Keywords: Neuroethics, neuroscience, ethics, neuroimaging, enhancement, responsibility, privacy, free will, moral judgments, criminal responsibility

Katrina L. Sifferd JD, PhD, Assistant Professor of Philosophy, Elmhurst College, 190 N. Prospect, Elmhurst, IL, 60126, USA

Glossary Entries:

Neuroethics: The body of work exploring the ethical, legal, and social implications of neuroscience.

Neuroimaging: Technologies such as computed tomography (CT), anatomical and functional magnetic resonance imaging (aMRI and fMRI), and positron emission tomography (PET), which can display the structural and functional neurobiological bases of our cognitive capacities and psychological states.

Magnetic Resonance Imaging (MRI): an imaging technique used to visualize detailed internal structure and function of the body. MRIs use a powerful magnetic field to align the nuclear magnetization of atoms in the body. Radio frequency (RF) fields are then used to produce a signal that can construct an image of the internal structure of the body. Functional MRIs (fMRIs) measure signal changes in the brain due to changing neural activity by detecting increases and decreases in blood flow.

Electroencephalography (EEG): The recording of the electrical activity along the scalp produced by the firing of neurons within the brain. These electrical voltages, often called brain waves, vary as a function of time and type of brain activity (e.g., the EEG of a person in a coma will look different to the EEG of a normal, awake adult).

Position Emission Tomography (PET): An imaging technique that produces an image of functional structure in the body. The technique uses a tracer which is introduced into the subject's body whose presence is then detected and reconstructed into a three-dimensional image.

Dualism: The view that humans are made out of two types of stuff, physical stuff and a non-physical mind or soul.

Physicalism: The view that humans are made out of only one type of stuff, physical stuff (and thus there is no immaterial soul or immaterial mental states).

Libertarianism: The view that free will exists only where an act is uncaused by prior physical effects, but is not random or uncaused. This position usually entails the idea that an immaterial self or soul is the causal agent in human action.

Compatibilism: The view that human action can be understood to be free even if it is fully determined by a physical-causal chain of prior events, if the act is linked in a particular way to a person's desires or psychological states.

Brain Fingerprinting: A technique which may be able to detect whether a person recognizes a picture, word or phrase. Brain fingerprinting uses EEG readings in an attempt to identify changes in brain activity when a subject is presented with familiar information. These EEG readings are then compared to the subject's verbal report regarding what he knows.

Synopsis:

Neuroethics is the body of work exploring the ethical, legal, and social implications of neuroscience. This work can be separated into two rough categories. The neuroscience of ethics concerns a neuroscientific understanding of the brain processes that underpin moral judgment and behavior. The ethics of neuroscience, on the other hand, includes the potential impact advances in neuroscience may have on social, moral and philosophical ideas and institutions, as well as the ethical principles that should guide brain research, treatment of brain disease, and cognitive enhancement. Central to the questions posed in neuroethics is the way in which neuroscience might impact our sense of self and personal responsibility, and our understanding the structure of moral judgments.

Body:

The term "neuroethics" was coined in 2003 by political journalist and New York Times columnist William Safire. At a conference convened by Safire and the Dana Foundation, of which Safire was the Chairman, Safire defined neuroethics as "the field of philosophy that discusses the rights and wrongs of the treatment of, or enhancement of, the human brain." He then charged 150 neuroscientists, bioethics scholars, policy-makers, and journalists with the task of exploring the ethical, legal, and social implications of neuroscience.

Safire's definition has since been expanded to include both the ethical implications of neuroscientific research – the ethics of neuroscience – but also the neuroscience of ethics. According to Adina Roskies, the neuroscience of ethics is concerned with a neuroscientific understanding of the brain processes that underpin moral judgment and behavior. The ethics of

neuroscience, on the other hand, includes the potential impact advances in neuroscience may have on social, moral and philosophical ideas and institutions, as well as the ethical principles that should guide brain research, treatment of brain disease, and cognitive enhancement.

This entry will discuss these different aspects of neuroethics, with a special focus on the way in which neuroscience might impact our sense of self and personal responsibility, a concern that cuts across these categories. For example, I will discuss whether advancing knowledge of brain states or processes undermine common notions of free will and responsibility. I will also examine whether certain treatments of brain abnormality are ethical (is it acceptable to irreversibly 'cure' pedophilia or obsessive/compulsive disorder?), a discussion that falls squarely under the category of the ethics of neuroscience. Finally, I'll discuss whether the neuroscience of ethics can provide insight into who should be deemed criminally responsible via a neuroscientific analysis of intentional action.

Neuroscience, the Self, Free Will, and Responsibility

Neuroscience is a scientific discipline aimed at understanding how human behavior and capacities are related to brain structure and systems. Neuroscience includes neurophysiology, neuropsychology, and neuroanatomy, and utilizes techniques such as Electroencephalography (EEG) as well as brain imagining technologies such as computed tomography (CT), anatomical and functional magnetic resonance imaging (aMRI and fMRI), and positron emission tomography (PET). These technologies can display the structural and functional neurobiological bases of our cognitive capacities and psychological states and processes, such as decision-making, desires, memories and emotions.

Other disciplines that treat human beings as the objects of scientific inquiry include anatomy, biology and chemistry. However, as neuroscience focuses upon the operations of the nervous system, which is the locale of capacities closely associated with personality, character, and the self. Thus neuroscientific findings are thought to have deep philosophical implications with regard to who we are (and possibly, what we are).

So, who are we, according to neuroscience? Neuroscience describes the brain as a causal machine. This means the brain goes from state to state as a function of prior, or antecedent, conditions: if the antecedent conditions were different, the result would be different; if the antecedent conditions were the same, the result would be the same. For example, at the beginning of some of these functional state changes is perceptual (or sensory) input, and at the end of some series of changes is a behavioral output (action).

Somewhere in between the input (perception of the environment) and output (action) thoughts, preferences, memories, and decisions seem to reside. Neuroscience provides us with new means

to examine the way in which these mental states and processes are instantiated in the brain by allowing scientists to observe the brain undergoing processes in response to stimuli via functional MRI (fMRI) machines. Data from an fMRI scanner shows what sites in a subject's brain are active (indicated by an increase of the blood flow signal) during mental activity, such as the experience of particular conscious mental states. Additionally, anatomical MRI and PET scans allow scientists to view the structure of the brain, and thus allow for identification of abnormalities in brain tissue. It may then be postulated that such abnormalities impact brain function in particular ways.

Neuroscience is thus thought by some to comment upon three age-old philosophical problems: first, what is the nature of the human mind? Second, how is the brain related to human action? And third, for which acts may a person be held responsible?

Neuroscience and the mind

Long before neuroscience informed our ideas of how the brain operates, many philosophers and scientists were already convinced that the position of dualism – that humans are made out of two types of stuff, a physical part and non-physical mind or soul – was false. Many feel there is little evidence in support of an immaterial mind, especially if one thinks this immaterial mind has causal effects. Neuroscience adds to the already substantial body of evidence that brain states and processes either enable or constitute our psychological states and processes, including thoughts, memories, and even higher mental functions such as consciousness and reasoning.

The tools of neuroscience thus provide further evidence that the brain constitutes the mind, and may even eventually describe which parts of the brain constitute which mental states or processes. However, neuroscience does not necessarily rule out the possibility of a non-physical aspect of the mind. Some think that conscious states, such as pain or love, are just what it feels like to undergo certain neural processes. Frank Jackson, however, has argued that the conscious feel of a mental state, such as the perception of the color red, could be an immaterial "add-on" to a physical brain state. Jackson notes, however, that in this case it is most likely that such immaterial states, which he calls "qualia", are epiphenomenal: that is, they have no physical causal effects. This means that it would be the neural instantiation of the state of pain or thirst that causes behavioral effects, not the immaterial conscious feel.

Although neuroscience may not definitively rule out immaterial qualia, it does seem to indicate that the cognitive capacities we consider most representative of the 'self', such as decision-making capacities and modulation of emotions, can be functionally identified as discrete states or processes within the brain. fMRI studies have allowed us to isolate and locate within the brain important cognitive capacities by identifying stored representations (memories) as functionally dissociated from the processes that operate upon them (decision-making). Neuroscience indicates that the frontal lobes do not contain many of our mental representations or memories, and that they instead reside in the posterior cortical regions, in the temporal and parietal lobes. (An

exception may be motor representations, and some representations involved in procedural or working memory.) The prefrontal lobes primarily contain the thought processes, often called executive processes by neuroscientists, which monitor and manipulate representations.

As we shall see, neuroscience's increasing ability to provide a structural and functional description of the brain has far-reaching implications. First, it could allow us to designate a portion of the brain as particularly important – because it functions in a way crucial to decision-making, for example - and in need of protection against manipulation or even enhancement (see below). Second, neuroscience may be able to not only identify brain abnormalities, but to explain why certain disorders of the brain are experienced as particularly harmful. Third, neuroscience could provide us with a better understanding of who is fully responsible for their actions, and who is not, via analysis of the function of part of the brain most closely associated with the self and decision-making.

Free Will

Philosophers have long questioned whether the everyday concept of free will – the libertarian version of free will where a free act is uncaused by prior physical effects, but is not random or uncaused – is true. However, because neuroscience advances the idea that the brain is a causal machine, it has provided further evidence that this traditional libertarian theory of free will is false. Neuroscience appears to support what scientifically-minded philosophers had long surmised: that there are no explanatory gaps in our brain processes wherein the causal chain is broken and libertarian free will steps in to cause action. Neuroscientific evidence, like a biological understanding of the conditions under which a neuron fires (or fails to fire), indicates that one need not appeal to an immaterial self or 'soul' to explain human behavior.

In particular, the work of Libet caused many outside of philosophy to think neuroscience meant human free will was an illusion. Libet conducted experiments where he had subjects make voluntary hand movements while he measured their brain activity using event-related potentials. Subjects were told to report the position of a black dot when they made the conscious decision to move their hands. Then this moment was compared to the time when the subject's brain appeared to change in relation to the desire to move the hand.

Libet discovered that the brain appeared to initiate the sequence resulting in the hand movement before the subject was consciously aware that he or she had made a decision to move: about 300 ms elapsed between the beginning of brain activity and the decision. Thus it seems our brains know we are going to make a decision, and perhaps maybe even what the decision is, before we do. (The "we" here is the conscious agent.)

Libet noted, however, there is still about 100 ms after the initiation of the act for the conscious mind to either endorse or veto the decision. Thus he argued that free will was not entirely an illusion, claiming that it consists in this veto power consciousness may have over a decision.

Regardless of how one interprets Libet's results, neuroscience does not force one to abandon the concept of free will. Many philosophers now accept the position of *compatibilism* – that human action can be understood to be free even if it is fully determined by a physical-causal chain of prior events. Walter Stace, for example, argues that free acts are "those whose immediate causes are psychological states in the agent", regardless of whether those states are physically instantiated and part of a larger causal chain. That is, it may be that your desire to get a drink was initiated by some physical state equal to thirst, and that your decision to get up and open your refrigerator was causally determined by this physical state. But because these states are "yours" the action should be considered free. By contrast, unfree acts are acts whose immediate cause is some state of affairs external to the agent, such as one holding a gun to your head.

Who is responsible?

Quite a few scholars have raised the alarm regarding the use of neuroscience in the criminal courts precisely because they feel that criminal responsibility requires libertarian free will. But as indicated above, even if neuroscience provides a new sort of explanation of the causes of behavior, this explanation does not necessarily constitute an excuse. The current criminal justice system can hold persons responsible even if human actions have full physical explanations: that is, even if a desire to kill is neuronally instantiated and has a complete causal history, the criminal justice system is designed to attribute responsibility when that desire leads to criminal harm.

Criminal responsibility requires that action be immediately caused in a certain way. If a harmful act is connected to the actor's desires or goals, and if the actor held certain beliefs about the harm that could result from the act, then he or she is criminally responsible. Thus, to be guilty of murder under the U.S. Model Penal Code, for example, one must have a *mens rea* (or a mental state) that includes: (1) the desire to perform an act that results in the death of another; and (2) performance of the act with the desire to kill, knowledge one will kill, or with reckless indifference to the chance that one may kill.

It is possible (although unlikely, at least given the current state of the science) that neuroscience may provide a complete physical causal description of the mental states required for criminal responsibility -- the desire to kill and the understanding that the act will result in a death. But even if this were to happen, it need not affect the categories of culpability (such as murder, rape, and theft) themselves. The breadth of category is a policy decision determined by legislatures and judges. Neuroscientific knowledge can thus be understood as a tool to help us better categorize defendants as guilty or not guilty based upon the existing categories of culpability and defense.

However, there is some chance that neuroscience will cause us to alter or abandon the categories we use to attribute criminal guilt. Some have suggested that the vast majority of offenders on

death row suffer from severe prefrontal lobe damage. What if it turned out that the vast majority of murderers had severely compromised executive function, such that, using the delicate tools of neuroscience, we determined they couldn't either control or understand their violent acts in the way that those without compromised brain function can? In other words, what if they didn't form desires or understand the nature or consequences of their acts in a "normal" way? Would punishment still be justified?

In the US, criminal punishment tends to be justified on the basis of one or several of the "principles of punishment" such as deterrence, incapacitation, or retribution. Neither retribution nor deterrence would seem to justify severe punishment of persons with brain function compromised in a way relevant to criminal responsibility (e.g. with regard to intentional or rational action). Such persons would not be deserving of retribution if they truly didn't understand the nature of the act they were performing or were unable to stop themselves from performing it. Similarly, an understanding of the act as criminal and the ability to stop oneself from performing the act is crucial to the effectiveness of deterrence.

However, just as we incapacitate those deemed legally insane or "guilty but mentally ill" despite the fact that they are ineligible for punishment, we might also decide to incapacitate violent offenders who lack the capacity to understand or prevent their criminal acts. Thus in this case, neuroscience might alter our theoretical handling of offenders with abnormal brain processes, but our actual treatment of the offenders – that they are incapacitated – may not change.

Brain Research, Treatment and Enhancement

In addition to its implications for traditional philosophical concepts of mind, will and responsibility, there are other ethical concerns raised by advances in neuroscientific research. One has to do with issues of privacy; another with the 'self' and the way in which this self might be altered or enhanced via neuroscience.

<u>Privacy</u>

Take, for example, the use of neuroscience as a "lie-detector". The first lie detector consisted in a polygraph test. A polygraph is an instrument that measures and records physiological indices such as blood pressure, respiration, pulse, and skin conductivity while the subject is asked and answers questions. Although many dispute the accuracy of polygraph tests, some argue that lying produces a physiological response that can be differentiated from true answers.

A recently developed lie-detector technique, however, uses neuroscience. "Brain fingerprinting" devices use EEG readings in an attempt to identify changes in brain activity when a subject is presented with familiar information. An EEG records the electrical activity along the surface of the scalp produced by the firing of neurons within the brain. These electrical voltages, often

called brain waves, vary as a function of time and type of brain activity (e.g., the EEG of a person in a coma will look different to the EEG of a normal, awake adult). In brain fingerprinting, changes in EEG readings are compared to the subject's verbal report regarding what he knows. The device is thus thought to be able to determine whether a criminal suspect is familiar with some aspect of a crime, and hence whether the suspect is telling the truth concerning their involvement. Such a device could also be used by potential employers, or even insurance companies, to determine if a job candidate or person seeking insurance was being honest about their past work experience, or any preexisting conditions.

Because brain fingerprinting detects brain activity directly, it is thought by some to be a more accurate lie detector than use of a polygraph. This is because it does not entail detection of emotional responses, which can be consciously manipulated by some subjects. However, there is controversy over whether brain fingerprinting returns consistently accurate results. In addition, brain fingerprinting can only indicate whether a subject is familiar with certain information; the technique cannot indicate how the subject gained this information (e.g. via being present at the crime scene, or reading about the crime in the newspaper).

There are other ways that neuroscience may provide access to previously private mental states. Newly developed quantum dot technology is being used to gather information in the brain at the level of the neuron. Nano-sized functional quantum dots can help build data-capture devices that are easy to use by neuroscientists. Many feel that nanotechnology in conjunction with neuroscience will eventually allow for targeted interactions between neurons, the cells responsible for signal transmission in the brain. Ultimately, it seems clear that nanotechnology will allow us to visualize and track functional responses in neurons, and this means we will be provided information about a person's thoughts remotely.

Several ethical issues are raised by the possibility of use of brain fingerprinting devices or functional quantum dots. Normally, an employer or a police officer must rely upon information voluntarily offered: regardless of what one is thinking, the information actually spoken is what counts, and most persons have control over the words they chose to speak. In so far as neuroscience allows others direct access to one's consciousness or mental states, it would appear to entail a serious violation of privacy. And in the cases mentioned above, the consequences of the violation are severe: one might not get a job or qualify for health insurance, or in the case of the criminal law, one could lose their liberty or life.

Enhancement

The possibility of brain enhancement via neuroscience brings forth other ethical worries. Pharmacological enhancement of brain function is already widely practiced via drugs such as selective serotonin reuptake inhibitors (SSRIs, prescribed for mood enhancement), Ritalin (for improved attention and concentration) and Aricept (for improved memory). Although originally conceived to make up for deficits in cognitive ability due to disorders such as depression and

attention deficit hyperactivity disorder, such drugs are now often used by those with highly functioning, and even "normal" cognition, such as those diagnosed as having a vulnerability to depression (but without any current depressive symptoms).

Enhancement may also be achieved by brain implants. Brain implants have been used to lessen the effects of several common disorders, such as Parkinson's and chronic pain. For example, in the case of Parkinson's, a technique called Deep Brain Stimulation (DBS) uses a brain implant to deliver electrical impulses to a region in the center of the brain that controls and coordinates movement. DBS has also been used in an attempt to stimulate various parts of the brain which appear dysfunctional in those who suffer from major depression and Tourette's syndrome. In the future, it is possible that DBS could be used to enhance cognitive function on those without psychological disorders.

Scientists are currently making progress using both carbon nanotubes and conventional silicone implants in the brain. Ultimately, such implants may be used to manipulate thought, possibly via transmission or implantation of thoughts. One possible use of this technology could be to allow soldiers to fight better under combat situations, or to allow athletes to better focus on their sports.

One ethical concern regarding brain enhancement is the idea that in order for humans to fully enjoy or appreciate life, they must experience both joy and pain. Overuse of SSRIs or other brain enhancing drugs might thus blunt a person's ability to enjoy the full range of human experience. Another worry has to do with the way in which enhancement will be distributed. It is likely that the wealthy and privileged will have more access to such enhancements, thus creating an even greater gap between socioeconomic classes. For example, a drug that assists in memory may mean that those who can afford the drug may have an advantage in the job market, and athletes with neuro-enhancements will be able to out-compete those without enhancement.

Yet another concern regarding neurological enhancement is that it is done in accordance with a normative idea of the optimal human being or life. However, there is much disagreement about the "good life" for human beings, and even capacities we might all agree are beneficial with regard to specific tasks could turn out to have negative outcomes overall. For example, there is some evidence that the average human level of forgetfulness actually helps us attend to the things that are important, instead of being distracted by unnecessary details. There is certainly no evidence that having better mental capacities will make human beings happier.

There is also the possibility of the use neuro-enhancement or manipulation in the criminal justice system. For example, a handful of US states have chemical castration laws, which means that at least some subset of their criminal offenders either must or may submit to chemical castration as a part of their sentence. Depro-Provera is the drug most often used for chemical castration. The chemical used is an analogue of the female hormone progesterone, which reduces the normal level of testosterone in a male, thus reducing sex-drive, and often diminishing seminal ejaculator fluid to zero.

Depro-Provera has to be continually administered, which requires non-incapacitated offenders to submit themselves to weekly shots. One might imagine neurological castration as a replacement to this drug regimen. Neurological castration might directly inhibit activity in certain parts of the brain (e.g., within the hypothalamus), by blocking connectivity between areas of the brain (e.g. between representations of children and sexual arousal). One might imagine that a nano-technological approach to castration may be more successful, and have far fewer side effects, than current methods.

Neurological castration could just be the beginning of neurological sentencing techniques. If it became possible to neurologically inhibit strong violent responses to stimuli, the state might offer offenders the chance to submit to this operation in exchange for a shortened or commuted sentence. Granted, at the moment this possibility is more fiction than science. However, given the success in drug interventions on aggressive behavior – for example, with tranquillizers and some anti-depressants – it is possible that neuroscience could discover a more targeted means of delivering the same result.

Neuroscience could thus potentially allow us to re-embrace the principle of rehabilitation by providing a means to directly change behavior. One might even wonder whether we could eventually come up with what some might call an "artificial conscience." Up to recently, persons who violate the law are allowed to remain the sort of person they are (even if that person was a pedophile), although the space within which they are allowed to be that person is limited to a jail or prison. And after an offender served their time, persons are released to continue to pursue their individual desires. However, neuroscience instead might allow us to permanently change a person's character, such that she no longer commits criminal acts.

Any neuroscientific manipulation or enhancement that changes a person's loci of decision-making represents nothing less than changing their identity. Most philosophers believe that one's psychological states – our beliefs, desires, memories and emotions – are a crucial component of one's identity. Our psychological states are the 'source' of our behavior: we act a certain way because we desire and believe certain things. If neuroscience changes what we want, what we know, or what we believe, it changes who we are.

With regard to criminal offenders, involuntary manipulation of psychological states via neuroscience would be a severe violation of human agency in that it would infringe upon the ability of a human being to choose, and be responsible for, his own acts. However, use of technology in this way raises serious ethical concerns even if offenders participated in such programs voluntarily. Would we really be giving offenders a choice with regard to neuro- or nano-technological alteration if they must choose between thirty years in prison and freedom after a "simple" operation? There are real worries that such an option would be no option at all.

A Neuroscientific Understanding of Moral Judgment and Action

Traditionally, philosophers such as Immanuel Kant argued that higher-level reasoning was crucial to moral judgment and action. However, it has become increasingly clear that moral judgments are not the sole product of reason or introspection isolated from emotion. Indeed, it turns out that emotional input is crucial to moral action and the judgments we make regarding whether others' actions are right or wrong.

Neuroscientific research attempting to isolate the regions involved in moral reasoning and behavior support this conclusion. Neuroimaging indicates that there is no particular "moral center" of the brain. Widely distributed brain areas have been found to be activated during moral reasoning tasks, including those regions associated with higher cognitive functions and those involved in emotional responses. There appears to be a complex interplay between reasoning and emotions in the formulation of a moral decision, indicating that there may not be an easy distinction between thinking and feeling when it comes to morality.

It is somewhat difficult to parse apart findings indicating the brain regions associated with moral judgments from those regions associated with behavior that we might judge immoral or "bad". This is because being able to determine that a situation requires a moral judgment, and treating a decision as a moral decision, may be important to moral behavior. If don't see my decision to steal my neighbor's TV, or not to give money to the Red Cross, as a moral judgment with ethical implications, then it seems I am less likely to make a decision others will approve.

Moral Judgments

Studies have found significant neurological differences between subjects making judgments on moral personal dilemmas, and those contemplating impersonal moral dilemmas and non-moral dilemmas. Moral dilemmas involving a personal component – where one is asked to consider whether they would personally perform some moral or immoral act – seem to activate the orbital, ventromedial and dorsolateral prefrontal cortex (PFC), the medial frontal gyrus, and possibly the posterior cingulate cortex, amygdala and inferior parietal lobe. These regions are associated with integration of emotions, mental imagery, and memories into decision-making, representation of rewards and punishments and control of risky or inappropriate behavior. In addition, the medial frontal gyrus, anterior and posterior cingulate cortex, and temporal areas (such as the superior temporal sulcus and temporal pole) are associated are associated with theory of mind, the ability human beings have to attribute mental states to others as a means to understand and predict their behavior.

Antonio Damasio and colleagues have found that patients with damage to the posterior ventromedial cortex have impaired ability to make moral judgments. However, their impairment was primarily emotional; they appeared to have an inability to generate and effectively use what Damasio calls "somatic markers": neural representations of body states that give potential bodily

actions a (sometimes unconscious) emotional feel making them more or less appealing. If you ask a subject without ventromedial damage to imagine running over another person in their car, they tend to have a negative emotional response in association with the mental image. However, although subjects with ventromedial damage could imagine performing the act and its consequences, and knew that society would judge such as act as wrong or criminal, they failed to have a normal emotional response to the mental image.

It seems the more personal a moral dilemma, the stronger one's potential emotional response. Greene and Haidt scanned subjects using an fMRI while they responded to a series of personal and impersonal moral dilemmas, as well as non-moral dilemmas. They determined a moral violation is personal if it is clear that the subject was hurting some particular other person, whereas it is impersonal if it didn't require the subject directly harm someone else. A classic philosophical thought experiment, for example, asks whether the subject would pull a switch to move a trolley car from one track, where five people are likely to be hit by the trolley and killed, to a second track, where only one person is going to be hit and killed. This is an impersonal moral dilemma, because it does not require that the subject's act directly cause harm, but only deflect harm from one party to another. On the other hand, when subjects are asked to push a fat man onto the tracks to derail the trolley to keep the five persons from being killed, the dilemma becomes personal, because it requires direct harm be caused by the subject to a person who would otherwise be safe.

Greene and Haidt found that responding to personal moral dilemmas, as compared with impersonal and non-moral dilemmas, produced increased activity in the medial frontal gyrus, posterior cingulated gyrus, and the angular gyrus. These areas are associated with social and emotional processing. Impersonal and non-moral dilemmas, on the other hand, produced increased activity in areas associated with working memory, such as the dorsolateral prefrontal areas. These findings again indicated the importance of emotional input in moral decision-making: subjects faced with a moral dilemma tended to be quick to moral condemnation but slow to approval of personal violations, apparently due to the salience of negative emotions, whereas approvals and disapprovals took equally long when subjects were faced with impersonal and non-moral dilemmas.

In sum, neuroscience indicates that many different capacities are important to moral reasoning and judgment, including normal decision-making and theory of mind capacities; however, the importance of emotional input may be particularly important to moral judgments (as opposed to many other types of judgments).

Moral (or Immoral) Behavior

As indicated above, there is significant overlap between the brain regions associated with moral judgment, and the regions associated with immoral or antisocial behavior. Regions common to both include the ventral and medial PFC, the amygdala and the angular gyrus/superior temporal

gyrus. One difference, however, is that persons who tend toward antisocial behavior show hippocampal and anterior cingulate impairment, whereas moral judgment studies fail to indicate such activation consistently in moral reasoning tasks. However, psychopaths also show fronto-temporal-limbic gray matter reductions (DeOlivera et al., 2008 NeuroImage) and reduction in connections between amygdala and OFC (Craig et al., 2009 Mol Psych).

Adrian Raine and collegues have argued that some of the neurological impairments found in antisocial individuals disrupt moral/emotional decision making, which in turn acts as a factor predisposing such individuals to antisocial behavior. Some antisocial personalities, and psychopaths, appear to show normal or excellent moral reasoning ability, but appear to fail to apply the outcome of such reasoning processes due to a lack of emotional input.

As an example, some argue that the ventromedial frontal cortex (VM cortex) mediates between the neural systems for arousal and emotion, and thus is important to moral behavior. Patients with VM cortex damage tend not to show deficiency in moral reasoning: that is, in their ability to identify potential moral conflicts and potential solutions to such conflicts. However, at least some VM patients *feel* differently about the identified solutions than those without VM damage. Going back to the trolley thought experiment, both VM patients and persons without VM damage realize that killing the one fat man will save five others lying in wait on the other trolley track. However, persons without VM damage tend to judge it wrong to kill him. Persons with VM damage, however, are more likely to judge that pushing the fat man onto the tracks is the right thing to do.

Not surprisingly the difference between the two judgments seems to be emotional: as indicated above, many think psychopaths act immorally partially due to a lack of normal affect. Thus, persons with VM damage may be more likely to engage in antisocial or immoral behavior, precisely because they do not feel badly about such action. In conclusion, many argue that it is predominately the failure to *feel* of what is moral, rather than the knowing of what is moral, that leads many to antisocial or immoral action.

Neuroscientific Evidence and the Criminal Law

From the perspective of the criminal justice system, it is difficult to understand what the conclusion that many antisocial persons lack emotional insight should mean. To the extent that neuroscience has thus far been used by the criminal law, it has been introduced to help the court determine whether a defendant had either minimal capacity to form the intent to commit a crime, or if the defendant was capable of forming the highest level of intent (e.g. the intent to kill in the first degree or "purposely", for which the defendant might qualify for capital punishment). In this way neuroscience has been understood to be a rough tool for commenting upon the basic decision making capacities necessary to understand the nature and quality of one's actions in a normal way.

For example, in 2003, Burns and Swerdlow reported a case of a 40-year-old schoolteacher in an otherwise normal state of health who developed an increasing interest in pornography, including child pornography. The patient also began soliciting prostitution at "massage parlors," which he had not previously done. Magnetic resonance imaging revealed a large tumor in his right orbitofrontal lobe. The patient's symptoms disappeared after the tumor was removed. The tumor in the frontal lobes had impacted several important higher cognitive functions (executive processes), including inhibition (denial of immediate gratification), the moderation of emotional responses, and consideration of the potential future consequences of actions. Burns and Swerdlow hypothesized that the tumor had not given the patient new goals or desires, but had rather impacted the patient's ability to conform his behavior to societal norms and laws.

The patient was arrested for sexual assault of his stepdaughter. If the case had gone to trial, it would seem that identification of the man's tumor would seem to be extremely helpful to an accurate understanding and categorization of his behavior. Indeed, it would appear to be unjust to deny a judge or jury a chance to consider the neuroscientific evidence of the presence of a tumor in the man's orbitofrontal lobes as evidence of diminished legal capacity.

However, this specific sort of use of neuroscientific evidence by the criminal law is quite different from imagining the implications of a finding that some high percentage of antisocial behavior is linked to brain abnormalities. In the above case, the tumor in the brain of the patient described by Burns and Swerdlow led him to behavior that was abnormal for the patient (and disappeared after the tumor was removed), indicating that less responsibility for an act linked to the tumor is appropriate. However, as stated above, if all murders have brain certain common brain abnormalities, does that mean we have a better indicator of the sort of person we feel is dangerous and want to punish, or does it mean such persons are less culpable? This question does not appear to have an empirical answer, and seems to fall instead within the realm of public policy.

(Greene and Haidt 2002) (Stace 1952; Smart 1959; Smith 1998; Libet 2002; Papineau 2002; Roskies 2002; Burns 2003; Casebeer and Churchland 2003; Dennett 2003; Churchland 2006; Illes 2006; Morse 2006; Raine and Yang 2006; Roskies 2006; Silva 2006; Glannon 2007)

Further Reading:

Burns, J. a. Swerdlow (2003). "Right Orbitofrontal Tumor With Pedophilia Symptom and Constructional Apraxia Sign " Archives of Neurology.

Casebeer, W. D. and P. S. Churchland (2003). "The neural mechanisms of moral cognition: A multiple-aspect approach to moral judgment and decision-making." Biology and Philosophy **18**(1).

Churchland, P. S. (2006). Moral decision-making and the brain. Neuroethics: Defining the issues in theory, practice and policy. J. Illes. Oxford, Oxford University Press**:** 3-16.

Dennett, D. (2003). Freedom Evolves, Penguin Books.

Gazzaniga M., (2008), *The Law and Neuroscience*, Neuron, 60: 412-415.

Glannon, W., Ed. (2007). Defining Right and Wrong in Brain Science: Essential Readings in Neuroethics. The Dana Foundation Series on Neuroethics. New York, Dana Press.

Greene, J. and J. Haidt (2002). "How (and where) does moral judgment work?" TRENDS in Cognitive Sciences **6**(12): 517-523.

Illes, J., Ed. (2006). Neuroethics: Defining the issues in theory, practice and policy. Oxford, Oxford University Press.

Libet, B. (2002). "The Timing of Mental Events: Libet's Experimental Findings and Their Implications." Consciousness and Cognition **11**(2): 291-299.

Morse, S. J. (2006). Moral and legal responsibility and the new neuroscience. Neuroethics : defining the issues in theory, practice, and policy. J. Illes. Oxford ; New York, Oxford University Press**:** 33, 50 p.

Papineau, D. (2002). Thinking About Consciousness. Oxford, Oxford University Press.

Raine, A. and Y. Yang (2006). "Neural foundations to moral reasoning and antisocial behavior." SCAN(1): 203-213.

Roskies, A. L. (2002). "Neuroethics for the new millenium." Neuron **35**: 21-3.

Roskies, A. L. (2006). A case study of neuroethics: the nature of moral judgment. Neuroethics: Defining issues in theory, practice and policy. J. Illes. Oxford, Oxford University Press**:** 17-32.

Silva, G. (2006). "Neuroscience Nanotechnology: progress, opportunities and challenges." Nature Neuroscience **7**: 65-74.

Smart, J. (1959). "Sensations and Brain Processes." Philosophical Review **66**: 141-156.

Smith, K. L. (1998). "Making Pedophiles Take Their Medicine: California's Chemical Castration Law." The Buffalo Public Interest Law Journal: 1-42.

Stace, W. T. (1952). The Problem of Free Will. Religion and the Modern Mind, Haper Collins 248-258.

Websites:

The Neuroethics Society

http://www.neuroethicssociety.org/

The Center for Neuroscience and Society at the University of Pennsylvania

http://www.neuroethics.upenn.edu/

The MacArthur Foundation's Law and Neuroscience Project

http://www.lawandneuroscienceproject.org/

The National Core for Neuroethics at the University of British Columbia

http://neuroethics.ubc.ca/National_Core_for_Neuroethics/Home.html

Author Bio:

Katrina Sifferd is the author of various articles and book chapters, including "In Defense of the Use of Commonsense Psychology in the Criminal Law," in the journal Law and Philosophy 25: 571-612 (2006); 'Can Baird's View of Adolescent Morality Inform Adolescent Criminal Justice Policy?' in Moral Psychology Vol. 3: The Neuroscience of Morality, Walter Sinnott-Armstrong ed., (MIT Press 2007); "Nanotechnology and the Attribution of Responsibility", in Nanotechnology Law & Business 5: 177-189 (2008); and (forthcoming) "Changing the Criminal Character: Nanotechnology and Criminal Punishment", in The Social Scale: Nanotechnology and the Weight of Justice (MIT Press).

Katrina also holds a J.D. from DePaul University College of Law. Before teaching at Elmhurst College she held the Rockefeller Fellowship in Law and Public Policy and a Visiting Professorship in philosophy at Dartmouth College. Prior to earning her Ph.D., Katrina worked as a senior research analyst on projects for the National Institute of Justice.