A JUDGE'S GUIDE TO NEUROSCIENCE: A CONCISE INTRODUCTION

Acknowledgments

This publication grew out of the work of the Law and Neuroscience Project, supported by a generous grant from the John D. and Catherine T. MacArthur Foundation. The Project, consisting of a wide variety of leading scientists, philosophers, and legal academics, including judges, began its work in June 2007 and will conclude the first phase of research and publication next year. We thank Andrew S. Mansfield for his work in editing this collection which was developed by Michael S. Gazzaniga and the Honorable Jed S. Rakoff, United States District Judge for the Southern District of New York.

Contributors

Floyd E. Bloom, M.D.

Floyd Bloom is Professor Emeritus at Scripps Research Department of Neuropharmacology and was editor-in-chief of Science Magazine for 5 years. He is a member of the National Academy of Sciences, the Institute of Medicine as well as the American Academy of Arts and Sciences. He has received numerous awards, including the Pasarow Award in Neuropsychiatry, and the Sarnat Award for Mental Health Research.

Howard L. Fields, M.D., Ph.D.

Howard Fields is Professor, Neurology and Physiology at the Ernest Gallo Clinic and Research Center, University of California, San Francisco and is a member of the Institute of Medicine of the National Academies and the American Academy of Arts and Sciences. He is also Director of the Wheeler Center for the Neurobiology of Addiction. He studies motivational systems and uses an experimental approach in animals to analyze the circuits in the brain that mediate pain and reward.

Michael S. Gazzaniga, Ph.D.

Michael Gazzaniga is a Professor of Psychology at University of California, Santa Barbara, Director of the SAGE Center for the Study of the Mind, and Director of The Law and Neuroscience Project. He is a member of the Institute of Medicine of the National Academies and of American Academy of Arts and Sciences. Among other books, he recently served as editor-in chief of The Cognitive Neurosciences IV, from MIT Press, which features the work of nearly 200 scientists and is a sourcebook for the field.

Scott T. Grafton, M.D.

Scott Grafton is a Professor of Psychology at the University of California, Santa Barbara and Director of the UCSB Brain Imaging Center. He is an expert in nuclear medicine, functional brain imaging and in behavioral neurology. His research involves elucidating the cognitive architecture that underlies action representation. He has provided expert testimony in numerous medical-legal cases over the past 20 years. He is currently serving as Co-Director of the Institute for Collaborative Technologies which is a multi university project base at USCB.

Kent Kiehl, Ph.D.

Kent Kiehl is an Associate Professor, Departments of Psychology and Neuroscience, University of New Mexico and the Director of Mobile Imaging Core and Clinical Cognitive Neuroscience, Mind Research Network, Albuquerque, New Mexico. His research focuses on the clinical neuroscience of major mental illnesses, with special focus on criminal psychopathy, substance abuse, and psychotic disorders (*i.e.*, schizophrenia). In association with The MIND Research Network, he uses noninvasive techniques for measuring brain function, including event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI).

Helen Mayberg, M.D.

Helen Mayberg is a Professor of Neurology and Psychiatry at Emory University School of Medicine, Atlanta, Georgia and the Dorothy C. Fuqua Chair of Psychiatric Neuroimaging and Therapeutics. Her research concerns the characterization of neural systems mediating mood and emotions in health and disease using functional neuroimaging. Her primary goal is to define brain mechanisms underlying major depression, with an emphasis on development of algorithms that will discriminate patient subgroups, optimize treatment selection, and provide markers of disease vulnerability. Dr. Mayberg has testified as an expert witness in numerous trials concerning the introduction and use of brain imaging evidence in legal proceedings.

Read Montague, Ph.D.

Read Montague is a Professor of Neuroscience at Baylor College of Medicine, where he also serves as Director of the College's Human Neuroimaging Laboratory. His work focuses on computational neuroscience; the connection between the physical mechanisms present in real neural tissue and the computational functions that these mechanisms embody. This work concerns decision-making and free will. His lab is a member of The Computational Psychiatry Unit (CPU), a unit dedicated to understanding the computational connections between biological mechanisms and psychiatric illness.

Louis J. Ptáček, M.D.

Louis Ptáček is Director of the Division of Neurogenetics at the University of California, San Francisco School of Medicine, where he holds the John C. Coleman Distinguished Professorship in Neurodegenerative Diseases and is a member of the Institute of Medicine and the American Academy of Sciences. He is also an investigator with Howard Hughes Medical Institute. Dr. Ptáček's research focuses on genetic diseases of muscle, heart and brain, and hereditary variation of human sleep behavior. His studies are leading to insights about the molecular basis of these human phenotypes and to similarities and differences between human physiology and that of other organisms.

Marcus Raichle, M.D.

Marcus Raichle is a Professor of Radiology, Neurology, Neurobiology and Biomedical Engineering at Washington University in St. Louis. He is a member of the National Academy of Sciences, the Institute of Medicine, and the American Academy of Arts and Sciences and is the recipient of dozens of awards. He has been instrumental in the study of human brain function through the development and use of positron emission tomography (PET) and functional magnetic resonance imaging (fMRI).

Adina Roskies, Ph.D.

Adina Roskies is an Associate Professor of Philosophy at Dartmouth College. She is trained in both philosophy and neuroscience. After obtaining a Ph.D. in neuroscience and cognitive science she served, among many other positions and fellowships, as Senior Editor of the neuroscience journal Neuron. She then obtained a second Ph.D. in philosophy from MIT. Her philosophical research interests include philosophy of mind, philosophy of science, and

ethics. She was a member of the McDonnell Project in Neurophilosophy, a working group aiming to integrate philosophical thought with neurobiological research.

Anthony Wagner, Ph.D.

Anthony Wagner is a Professor of Psychology at Stanford University. His research includes examination of the cognitive neuroscience of memory and cognitive/executive control in young and older adults. He also investigates memory retrieval mechanisms and interactions between forms of memory, cognitive control, aging, and the organization of the prefrontal cortex and medial temporal lobe.

Table of Contents

Introduction

Jed S. Rakoff, Judge, United States District Court, Southern District of New York

It is part of the human condition to think about how we think and to attach consequences to our conclusions. In law, for example, whether a person goes to prison (and for how long), and whether a person is liable for damages (and for how much), are the product not only of what action the person took but also of what the person's state of mind was when he or she so acted. Yet few of us are very good mind-readers, and the law has struggled both to define relevant states of mind and to devise ways of perceiving them.

In recent years, lawyers and judges have gradually become aware that science is developing new tools to help discover how the brain works. They hear that there have been major advances in something called "neuroscience," that this progress is somehow related to something called "brain scans" or "fMRI's," and that it is changing how we think about how we think. They are not necessarily aware of how it might have an impact on the law, or, indeed, already has.

Over the past three years, the MacArthur Foundation's Law and Neuroscience Project (on whose Governance Board I have had the privilege of serving) has sought to address the nascent interaction between law and neuroscience on many fronts. Toward the end of the first phase of this Project, in the winter of 2010, Mike Gazzaniga, the Director of the Project, asked me what questions judges might have about law and neuroscience that those associated with the Project might help to answer. In response, I asked a sample of my colleagues on the federal bench to identify the ten or so questions they would most like answered concerning law and neuroscience. To an amazing extent, the responders to my not-very-rigorous survey asked the same questions, and it is those questions that this little pamphlet attempts to answer. The Law and Neuroscience Project is also developing more extensive materials for judges that are intended to be of service as neuroscience enters the courtroom. This pamphlet is more in the nature of an introduction. However, the very breadth of the questions it undertakes to address reflects the growing perception among judges that neuroscience has the potential to be of great use, and a challenge, to many aspects of the law. If this little pamphlet can serve to clarify that perception and help meet that challenge, it will have served its purpose.

What is cognitive neuroscience?

Michael S. Gazzaniga, Ph.D.

What is cognitive neuroscience? Since George A. Miller and I coined the term ourselves, you would think it would be easy for me to define. Yet Leon Festinger, the great psychologist, always warned his students not to fall into the "trap of premature precision." That advice was taken to heart during the early years in this field. There were so many considerations and so much to learn. This is still the case. At the same time, some 30 years into the discipline, ideas about what we are doing are finally beginning to emerge.

The quick answer to the question is cognitive neuroscience is the field of scientific endeavor that is trying to understand how the brain enables the mind. Depending how successful we are, information gained from these studies would be of great importance to the legal system.

The history of scientists trying to understand how the brain works is long and distinguished. Ivan Pavlov, positing the reflex arc, and Paul Broca, undertaking brain localization studies, are but two in a long series of scientists contributing to our current understanding of how the nervous system is organized. Other giants in the field, including Brenda Milner at the Montreal Neurological Institute, began to show how lesions localized to specific brain areas in humans produced particular kinds of memory and cognitive deficits.

Animal models were developed to deepen these clinical insights by researchers such as Mortimer Mishkin at the National Institutes of Mental Health. My own earlier work with Roger W. Sperry on the so called "split-brain" patients, which revealed independent conscious systems could reside in one brain, cried out for a field of human neuroscience.

The field of cognitive science emerged at Harvard in the late 1950s and spread quickly. Fueled by insights into human language by the linguist Noam Chomsky, George Miller began to study its implications for psychological structure in a new experimental field called "psycholinguistics." Traditional measures of an experimental psychology lab were initially used such as measuring reaction times to making various kinds of judgments.

Still, in the early 1980s, the thinking about how complex mental representations were instantiated in the nervous system was fairly primitive. This was true even though cognitive scientists were developing highly sophisticated ideas about mental structures.

However, a new enthusiasm was driving progress in the field. Experimental psychologists felt liberated from strict behavioral views, discarding the idea that everything could be explained in stimulus-response or behavioral terms. The view that there were cognitive structures that had their own dynamics and laws took over psychology.

When the field of cognitive neuroscience took off, there was little or no neuroscience associated with cognitive science. Neuroscience didn't know such things could be studied and cognitive science didn't have any biological constraints placed on its rich repertoire of theories.

The 1980s saw several changes in the scientific landscape. Established biologically minded psychologists such as Michael Posner, Stephen Kosslyn, and Steven Hillyard among others, began to ask how their models of human attention or human mental imagery might be better understood by seeing how it was affected by brain lesions.

There was a belief that any general human mental capacity, such as imagery, could be modeled by seeing how various sub-components interacted to produce the whole. These sub-components were called modules and the hope was that particular brain areas would be identified that matched up with the models built by the cognitive psychologists.

Around this time, as the two fields of neuroscience and cognitive science were coming together, a major event occurred. Human brain imaging suddenly appeared on the scene, and with it, the ability to study the human brain in action. Everything from basic issues in perception to higher order mental activities, were fair game for study. It wasn't only a question of observing brain areas activated under specific cognitive conditions. It was revealing which systems are involved in particular mental activities, such as mental imagery. Years of debate ensued on whether visual images in the mind relying on the actual brain systems known to be essential for vision could now be studied and resolved.

Over the past five years, the field has further expanded with methodologies of brain imaging directed at capturing mental states which are of particular interest to the law. Could we now examine the condition of the major mental states associated with *mens rea* as to their normal functioning? Can brain imaging technologies enlighten the court on whether or not someone has the capacity to form intentions or merely is reactive to circumstances?

Is there a neuroscience that can determine if someone does or doesn't have the capacity to judge the wrongfulness of an act because of an abnormal emotional nexus? Could this result in someone taking undue risks, which is to say act recklessly? Could someone whose brain biases him toward joy seeking in life behave negligently as a result?

These issues are now open for investigation. They are being studied and preliminary knowledge is being accumulated on all these topics.

For example, there are specific brain areas associated with the formation of intentions to act. It is easy to see how someone with brain damage in these areas might well be judged not able to form intentions and in some sense, lack a guilty mind. There are other lesions that affect how one feels about an act and that could well affect their tendency to act inappropriately. The list is long and rich and intriguing.

The following articles by various experts in the field of neuroscience offer views on a specific set of questions generated by judges. These are questions they would like to have answered, or at least receive guidance on, as they deal with issues of today.

Even though much of the impact of neuroscience will be down the road when methods yield more specific results with little or no room for mixed interpretations, the issue is upon us now for establishing a clear framework for conceptualizing how the science ultimately will be used in the courts.

General References

Gazzaniga, M. S., Ivry, R. B., & Mangun, G.R. (2008). *Cognitive Neuroscience: The Biology of Mind*. New York: Norton & Co.

Purves, D., Brannon, E. M., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., & Woldorff, M. (2008). *Principles of Cognitive Neuroscience*. Sunderland, MA: Sinauer Associates.

What is an fMRI?

Marcus Raichle, M.D.

During the past 30 years the field of cognitive neuroscience has emerged as an important growth area in the study of the human brain in health and disease. Cognitive neuroscience combines experimental strategies of psychology with various techniques to examine how brain function supports mental activities. Two techniques are now at that forefront of research in humans: positron emission tomography (PET) and magnetic resonance imaging (MRI).

This revolution in the study of the human brain began with the invention of X-ray computed tomography followed by PET and then MRI (Raichle 2009). Because of the dominant role now played by MRI in the study of human brain-behavior relationships and, therefore, the likelihood that it will be knocking on the courtroom door, I will focus on it exclusively.

Magnet Resonance Imaging (MRI)

Magnetic resonance imaging or MRI is based upon a set of physical principles associated with the behavior of atoms in water in a magnetic field. When placed in a strong magnet field, these water atoms (usually referred to as protons) behave like tiny bar magnets by lining up in parallel with the magnet field. When these protons are disturbed from their equilibrium state by radio frequency pulses that are delivered in the scanner, a voltage is induced in a receiver coil placed outside of the subject's head that can be characterized by its change in magnitude over time. Because these time dependent changes in voltage are a function of the local environment of the protons, many important deductions can be made about the tissue being examined.

When a strategy was developed in 1973 to create three-dimensional images of the human brain from these signals, the interest both clinically and scientifically was immediate. The technique behind these images has a unique sensitivity to the soft tissues of the body, such as the brain. Thus, MRI affords access to both detailed anatomy and function.

Functional Magnetic Resonance Imaging (fMRI)

Functional MRI, or fMRI as it is now called, emerged from three important discoveries. Two are old and one is more recent. It was in 1937 that the late Linus Pauling and his colleague

Charles Coryell noted that when hemoglobin (the red pigment in red blood cells that carries oxygen from the lungs to the body's organs) releases its oxygen to an organ, such as the brain, it becomes paramagnetic. Being paramagnetic means that hemoglobin will disrupt a magnetic field through which it passes, a prime example being the magnet field inside an MRI scanner. Thus, while arteries whose red blood cells carry a full load of oxygen do not disturb the magnetic field of the MRI scanner, veins do because they are carrying less oxygen and are therefore paramagnetic.

The second discovery important for understanding fMRI is that changes in brain activity are accompanied by local changes in brain blood flow. This was first noted by the great Italian scientist Angelo Mosso in 1881 and has been replicated countless times since. The critical point in connecting this observation with the fact that hemoglobin without oxygen is paramagnetic is the more recent discovery that when the brain increases its activity locally, blood flow increases more than needed to supply the increase in oxygen consumption. As a result, oxygen supply exceeds demand, hemoglobin loses less oxygen and the MRI signal increases locally.

Siegi Ogawa, David Tank and their colleagues working at the Bell Laboratories referred to this as the "blood oxygen level dependent" or BOLD signal, a name that has become synonymous with fMRI imaging ever since. Since the introduction of fMRI BOLD imaging, the growth of functional brain imaging has been nothing short of spectacular. A recent check on the number of scientific publications in which fMRI BOLD imaging was used revealed over twelve thousand publications since its introduction in 1992.

However, it is critically important to understand that functional brain imaging with fMRI is not like taking a picture with your iPhone. Success is not only the product of relevant physiology that can be imaged and the scanning devices that can accomplish this, but also the use of sophisticated strategies for the processing and statistical analysis of the image data.

Success is also dependent on the design of behavioral paradigms that approach human behavior in a principled and quantitative manner while accommodating the constraints of the imaging environment. I turn to these important issues next.

Image Processing

Stereotaxy

As images of task-induced changes in regional blood flow started to accumulate in the 1980s, an old problem resurfaced. How do you objectively relate functional imaging data to brain anatomy? This problem was neither new to functional brain imaging with PET nor previously unexplored.

The solution came in the form of a technique called stereotaxy, which was first developed by Horsley and Clarke for animal research in 1908, and much later applied to humans by neurosurgeons. Stereotaxy in humans is generally based on the assumption that all points in the brain of an individual have a predictable relationship to a set of landmarks. These landmarks can be easily identified by MRI, permitting an exact relationship to be established with a "standard" brain in one of the stereotaxic brain atlases.

While continuing to be modified and refined, this strategy has remained a central feature of all functional brain imaging in humans. Interesting recent advances have included the development of an ever more refined set of atlases by an international consortia of researchers to which all imaging data can be referred. There has been a tremendous advance in our ability to communicate these data in the scientific literature.

Image Averaging

The initial application of stereotaxy in functional brain imaging was to determine the location of activity changes in individual subjects. This approach worked rather nicely for robust responses that could be appreciated in individual difference images. However, other early experiments yielded data in which the responses were not robust and varied from subject to subject in location. The interpretation of the data was easily confused by what was termed image "noise." These data proved to be problematic and generated a great deal of concern in the functional imaging community.

In response to these problems and concerns, an effort was mounted to obtain averages of images across groups of subjects in a standard stereotaxic space. The wisdom of this effort was not universally embraced due to the fear that individual differences would simply be too great for success.

When the very first set of averaged blood flow images were produced, it was obvious to everyone that image noise was dramatically reduced and responses were crisp and clear. The processing of functional brain images had taken a major step forward.

Image averaging is now a key element in the processing of virtually all functional image data.

There is one very important and obvious assumption that is made when averaging is performed across a group of individuals: the way the brain functions during the task of interest must be essentially the same in all individuals for the averaging to work. The success of averaging groups of individuals is testimony to the fact that there are remarkable similarities at some level in the way individual human brains work.

Difficulty arises when making this same assumption in examining averaged data from a single individual when he or she performs the same task repeatedly. Experience has taught us that some tasks like reading common words aloud, a highly learned task, produces remarkably similar results within and across individuals no matter how many times the task is performed.

However, this is not the case when the task requires the subject to perform an operation that conflicts with the usual way of doing things. The best known example is asking individuals to generate a verb for common English nouns (*e.g.*, read "hammer," say "hit"). The reflexive response is to say "hammer." This must be inhibited so that the required response "hit" can be produced. The brain has very clever machinery to handle such a situation and its activity is readily seen with functional imaging.

However, and here is the important point, if this task is practiced for even a short period, the brain quickly converts it to a routine, reflexive operation which then, from the imaging perspective, looks just like reading the word aloud.

The most immediate judicial implication of the practice arises in the use of fMRI as a means of lie detection. The background of the subject and the questions that are asked are of critical importance in how the brain will respond. Lying for the first time in a mock situation is fundamentally different from a real life situation in which responses to questions can be anticipated and rehearsed. It is presently unclear how proponents of fMRI for lie detection propose to deal with this matter.

Statistical Analysis

With the above strategies in hand (*i.e.*, stereotaxic normalization and image averaging), investigators were suddenly confronted with yet another challenge: images containing enormous amounts of data. The specter of unacceptably high false discovery rates loomed large without an obvious remedy.

One obvious approach would have been to place independently determined regions of interest within difference images to test specific hypotheses about how the task under investigation was instantiated in the brain (*i.e.*, a traditional hypothesis testing). The major drawback to this approach was that it assumed the very knowledge one was seeking, namely, how the brain is organized! What was needed was a hypothesis generating approach.

Many of the details of how these uniquely challenging statistical questions were addressed are now of historical interest only. Statisticians, statistically minded neuroscientists, and others quickly found the problems inherent in the analysis of functional brain images stimulating and challenging.

From these important beginnings the approaches have become increasingly sophisticated, varied and powerful. It must, however, always be kept in mind that the validity of findings in any functional image data is critically dependent upon the statistical analysis strategy employed. Because of the sophisticated nature of the strategies now employed, of which there are many, an expert independent opinion is often needed when deciding upon the validity of the scientific evidence contained in functional image data.

The Behavioral Agenda

The study of human cognition with PET was aided greatly by the involvement of cognitive psychologists in the 1980s whose experimental designs for dissecting human behaviors using information-processing theory fit extremely well with emerging functional brain imaging strategies (Posner & Raichle, 1994). It may well have been the combination of cognitive psychology and systems neuroscience with brain imaging that lifted this work from a state of indifference and obscurity in the neuroscience community in the 1970s to its current role of prominence.

This strategy was based on a concept introduced by the Dutch physiologist Franciscus C. Donders in 1868. Donders proposed a general method to measure thought processes based

on a simple logic. He subtracted the time needed to respond to a light (say, by pressing a key) from the time needed to respond to a particular color of light. He found that discriminating color required about 50 milliseconds. In this way, Donders isolated and measured a mental process for the first time by subtracting a control state (*i.e.*, responding to a light) from a task state (i.e., discriminating the color of the light). This strategy was first fully implemented in the study of single word processing and has since been exploited with exponentially increasing sophistication in all aspects of functional imaging that is now addressing virtually all elements of human behavior in health and disease. One could easily see how the strategy might be employed in lie detection where neutral questions (the counter condition) are paired with crime-relevant questions. While simple in concept this approach has so far been difficult to implement and test.

Individual Differences

There is a natural tendency to extrapolate insights from functional imaging data presented in the scientific literature and even in the popular press to individual subjects. In evaluating the potential of brain imaging data to speak to the function or dysfunction of an individual subject, it must always be kept in mind that our understanding of human brain function from imaging is derived almost exclusively from averaging across groups of individuals. This has been necessary because of the usually poor statistical quality of individual subject data.

As a result, it is very challenging to predict from the averaged group data the exact nature of the function or dysfunction in individual subjects. Attempts to do this must be undertaken with caution. However, because of the obvious clinical importance of using functional brain imaging in the assessment of individuals, imaging of individual differences has become the focus of intense research in all aspects of imaging.

This research will likely lead to significant improvements in the future; notwithstanding, in the interim it is wise to take a cautious approach.

The Future

As we look to the future, even more dramatic changes clearly appear on the horizon. It may be too strong to suggest that we are facing a paradigm shift (Kuhn, 1996) but certainly some reorientation is taking place with how we understand brain function. While this

reorientation has received substantial stimulation from imaging work, it has its roots in more than a century of discussions about the nature of brain functions.

Since the nineteenth century and possibly longer, two perspectives on brain functions have existed (Raichle, 2010a; Raichle, 2010b). One view posits that the brain is primarily driven by external inputs; the other holds that the brain operates on its own, intrinsically, with sensory information interacting with rather than determining its operation.

While neither view is dominant today, the former clearly has motivated the majority of research at all levels of neuroscience, including that in cognitive neuroscience. This is not entirely surprising given the enormous success of experiments measuring brain responses to controlled stimuli.

From a cost-based perspective, however, intrinsic activity appears far more significant than evoked activity for overall brain function. Studies of the actual changes in energy consumption associated with evoked changes in brain activity have revealed that the additional energy required for such brain responses represents an extremely small percentage (~1.0%) of ongoing energy consumption (Raichle & Mintun, 2007). Furthermore, converging data suggest that 60-80% of the ongoing energy consumption reflects work associated with the input and output of neurons (Raichle & Mintun, 2006).

From this perspective it seems fair to conclude that a major fraction of the brain's functional activity is unaccounted for. What do we know about the organization of this activity from an imaging perspective?

In attempting to understand the brain's intrinsic activity, tremendous interest has focused on the "noise" in the fMRI BOLD signal. When resting quietly in an MRI scanner, the BOLD signal exhibits very slow fluctuations at about 0.1 Hz (*i.e.*, about one cycle every 10 seconds).

Because this was originally viewed as noise, it was conveniently eliminated by image averaging (see above) until it was discovered by Bahrat Biswal and colleagues that these fluctuations represent coherent activity within brain systems. Therefore, if one simply follows the time course of this activity in, say, the area of one cerebral hemisphere that controls motor output to the limbs on the opposite side of the body and asks what other areas of the brain are correlated with this spontaneous activity, quite amazingly it turns out that it is all the motor areas of the brain!

Using this approach scientists have now gone about mapping virtually all the human brain's major functional systems without ever having to ask subjects to perform a task (Fox & Raichle, 2006). This greatly expands the reach of fMRI to studies of subjects unable to perform tasks such as children and patients with various incapacitating diseases (Zhang & Raichle, 2010). This strategy also works well under general anesthesia and during sleep.

While this new approach to fMRI is not ready for the courtroom, it seems almost inevitable that it will eventually make it there, and probably sooner rather than later.

References

Raichle, M. E. (2009). A brief history of human brain mapping. *Trends in Neuroscience, 32*, 118-126.

Posner, M., & Raichle, M. (1994). *Images of Mind.* New York: W. H. Freeman and Company.

Kuhn, T. S. (1996). *Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Raichle, M. E. (2010a). Two views of brain function. *Trends in Cognitive Science*, *14*, 180-190.

Raichle, M.E. (2010b). The brain's dark energy. Scientific American, *302*, 20-33.

Raichle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience*, *29*, 449-476.

Fox, M. D., & Raichle, M. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews: Neuroscience, 8*, 700-711.

Zhang, D., & Raichle, M. E. (2010). Disease and the brain's dark energy. *Nature Reviews: Neurology*, *6*, 15-28.

Can neuroscience identify lies?

Anthony Wagner, Ph.D.

The advent of functional magnetic resonance imaging (fMRI) enables scientists to examine whether measurements of brain activity can indeed provide an accurate neuroscience-based approach to lie detection, in contrast to the polygraph, which indirectly assesses deception by way of measurement of nonspecific autonomic nervous system responses.

Development of a method capable of detecting lies with a high degree of sensitivity and specificity, while being resistant to counter measures, would affect the practice of forensics and the legal system itself.

A number of neuroscience techniques have been explored for this purpose, including electroencephalography (EEG) and positron emission tomography (PET). In recent years development efforts over the past decade have increasingly focused on fMRI. Two companies (No Lie MRI, Inc. and Cephos Corp.) offer commercial application of fMRI-based methods that they market as being able to distinguish truth from deception.

In this analysis, I discuss the state of the scientific literature and its bearing on whether fMRI-based neuroscience methods can identify lies.

Structure of Analysis

To date, there have been approximately 28 peer reviewed publications reporting unique fMRI or PET data sets that examine brain responses during putative "deception versus truth telling". This count excludes four papers in which previously published data were submitted to reanalysis and republished.

Twenty-one of these papers report data exclusively submitted to group-level analyses. These showed brain regions where a differential response between deception and truth conditions were identified using statistical procedures that displayed consistent effects across subjects.

Such studies have limited bearing on whether fMRI can detect lies at the individual-subject or the individual-question levels.

First I will briefly summarize what has been learned from these group-level studies. Then I will discuss the minority of studies that examined the ability of fMRI to detect whether an individual is lying. Because this second group of studies provides the only data that is

potentially relevant to determining whether fMRI can detect lies, I focus on critical methodological issues in these experiments along with the resulting data. As I show, fundamental methodological limitations ultimately render these studies uninformative.

It is my conclusion that there *are* no relevant published data that unambiguously answer whether fMRI-based neuroscience methods can detect lies at the individual-subject level.

Brain Responses during Deception and Truth: Group-level Studies

Functional imaging studies reveal brain responses by comparing neural activity differences during two conditions that are thought to differ along the factors of interest. Accordingly, the fMRI (and PET) studies that targeted correlates of deception in the brain compared activity during conditions in which subjects were being deceptive versus responding honestly.

In most studies, subjects made deceptive versus honest responses about past experiences, including whether they had viewed or currently possess a particular stimulus encountered within the context of the experiment (e.g., whether they had viewed or currently possess a particular playing card provided at the outset of the experiment).

In the majority of published studies, the experimenter instructed the subject to lie about a particular stimulus / past event and to tell the truth about others. Often a neutral baseline condition was included to serve as a common comparison condition for the lie and truth conditions.

The data was analyzed, which typically included two steps:

A) calculation of an individual subject's brain responses during the lie versus the truth conditions (or the comparison of each condition to the neutral baseline)

B) a group-level statistical analysis that revealed which brain regions differed between the conditions across subjects. In the twenty-one studies (nineteen fMRI / two PET) exclusively reporting group-level effects (i.e., the outcomes of analysis step "B"), findings in individual subjects (i.e., the outcomes of step "A") were not reported.

Five important points emerge when considering these group-level studies.

1. The vast majority of studies revealed a significant difference in activity, somewhere in the brain, during the putative lie versus truth conditions (or a greater difference between lie versus baseline than between truth versus baseline).

2. Across the studies, there is considerable variability in the particular brain regions showing a difference in activity between lie versus truth. No brain region was associated with deception versus truth telling in all studies. There are undoubtedly multiple reasons for this across-study variability, including differences in: (a) the types of lies examined (*e.g.*, Ganis *et al*., 2003); (b) the particular stimuli used and/or tasks performed (*e.g.*, modified versions of the guilty knowledge test using playing cards (see below for details) as compared to tasks probing autobiographical memories for events in the subject's recent past); (c) data acquisition procedures (*e.g.*, magnet field strength, acquired functional data resolution, number of trials per condition and number of subjects both of which impact a study's statistical power); and (d) the statistical procedures applied to the data including use of different statistical thresholds, and differences in whether the lie and truth conditions were directly compared or whether they were only indirectly compared with reference to the neutral baseline.

3. While there is notable variability in the brain regions observed across studies, there also is above-chance consistency. For example, a meta-analysis of data from twelve group-level studies that directly compared lie versus truth conditions (Christ *et al*., 2009) revealed a number of regions that were active across studies at an above-chance rate (though, again, no brain region was observed in every study). My lab has replicated the findings of this meta-analysis, drawing on data from a larger portion of the published literature.

4. Many of the regions consistently observed across studies – dorsolateral and ventrolateral prefrontal cortex, posterior parietal cortex, anterior insula, and anterior cingulate cortex – do not support mechanisms specific to lying, but rather support mechanisms that are engaged during a variety of cognitive tasks, including working memory and inhibitory control (Christ *et al*., 2009), as well as attentional orienting and memory for past events.

5. While anatomical consistency is present across studies, few experiments were specifically designed to examine why a given brain region was more active during the lie versus truth conditions (*e.g.*, do the effects, among numerous possibilities, reflect memory processes, attentional orienting to salient stimuli, the need to inhibit a competing response, or emotional reactions during instructed deception?). The vast majority of published

conclusions about the role of particular brain regions in deception are tentative, awaiting further directed experimentation.

Can Neuroscience Detect Lies?

Can fMRI-based neuroscience methods identify lies? From forensic and legal perspectives, this is a question that can only be addressed by determining the sensitivity and specificity of fMRI-based lie detection at the individual-subject and/or individual-question level. Studies that exclusively report group-level analyses do not provide any data that bear on this question.

There are 11 peer-reviewed papers that report an assessment of "lie" detection classification at the individual-subject level (including three reporting different analyses on the data set of Kozel *et al.*, 2005).

A total of nine unique data sets have been submitted to individual-subject analysis. Across these data sets, three types of tasks have been used: modified versions of the Guilty Knowledge Test (GKT) or Concealed Information Test (CIT); a "mock theft" paradigm where subjects give truthful versus deceptive answers about the location of money in a room or about which of two objects the subject took at the outset of the experiment; and a "mock sabotage" study involving the instructed destruction of property.

Studies using a Modified GKT or CIT

In four papers, data from three experiments using variants of a GKT or CIT were analyzed at the individual-subject level. In two GKT experiments (Langleben *et al.*, 2005; Monteleone *et al.*, 2009; with Davatzikos *et al.*, 2005 reanalyzing the Langleben data), the subject was initially presented an envelope that contained two items (*e.g.*, two playing cards); in the CIT experiment (Hakun *et al.*, 2008), the subject initially picked a number between 3 through 8.

Subsequently during scanning, the subject was to deny possession of one of the two items or deny having picked the number chosen (the "lie" trials) and acknowledge possession of the other item or deny having chosen the other non-chosen numbers (the "truth" trials). Other distractor items also appeared during scanning, to which the subject was to respond accurately; these distractors often served as a baseline for the lie and truth conditions.

In the work of Langleben and Davatzikos, analyses focused on whether lie and truth could be discriminated at the single-event level. That is, within a subject, was it possible to know

whether the subject was responding honestly versus deceptively on any given trial. The accuracy rates were 78% and 88%, respectively (with a 90% sensitivity and 86% specificity, in the latter study).

While potentially impressive, these data are rendered difficult to interpret because of a critical design limitation—the frequency of the motor response required on truth trials (a left button press) was much lower than that of the motor response required on lie trials (a right button press).

Because of this lurking variable, rather than reflecting the pure effects of deception, these data are likely to partly reveal brain differences associated with selecting an infrequent versus a frequent motor action. Potentially confirming the influence of this response variable on detection rates, Monteleone and colleagues analyzed data from a modified GKT where the response frequencies were matched for lie and truth trials, thus eliminating the statistical confound.

Using an analysis focused on identifying how many subjects showed greater brain responses to lie versus truth trials, Monteleone reported that 71% of subjects showed greater medial frontal lobe activity during the lie trials, with no subjects showing the reverse pattern. While this estimate does not bear on whether fMRI can discriminate between subjects who are lying from those who are not (because all subjects in Monteleone were instructed to lie), it is suggestive that fMRI-based methods may afford a better than average accuracy in detecting lies within an individual.

This conclusion, however, is greatly tempered by the report from Hakun and colleagues, who used a modified CIT. Hakun examined responses in brain regions previously shown to differentiate lie from truth trials at the group-level (in an independent group of subjects). Critically, in three out of three subjects, Hakun observed greater activation to the target stimulus relative to control stimuli when subjects were instructed to lie about having chosen the target and when subjects were simply passively viewing all the stimuli *(i.e.*, they were not instructed to lie).

This latter finding indicates that greater brain responses to a stimulus that a subject was instructed to lie about may not reflect processes related to deception, but rather may reflect an attentional orienting response to the stimulus because the experimental procedures rendered the "lie" stimulus more salient relative to the "truth" stimuli (*see also*, Gamer *et al.*, 2009).

Given the findings of Hakun *et al*., along with the response confound in the initial GKT experiments, I conclude that there are no unambiguous data from GKT/CIT paradigms that actually bear on whether fMRI can detect lies. [1]

Studies using a "Mock Crime" Paradigm

In two early experiments (Kozel *et al*., 2004a, 2004b), subjects were instructed to lie or tell the truth about the location of money in a room. While group level analyses in these studies revealed a number of brain regions in which activation was greater during lie versus truth responses, subsequent examination of the number of individual subjects who demonstrated a significant difference on lie versus truth trials in these brain regions revealed high variability. For example, Kozel *et al*. (2004b) observed that only 1 of 10 subjects showed a significant effect in the anterior cingulate cortex, and 5 of 10 subjects showed an effect in orbital frontal cortex. In 3 of the 10 subjects, none of the regions observed to differentiate lie versus truth at the group level showed a significant difference in these individuals.

The authors concluded that the "technique lacks good predictive power for individuals" (Kozel *et al*., 2004a) and that "The technique was not able to detect, within individuals, consistent activation patterns" (Kozel *et al*., 2004b).

Five subsequent papers (Kozel *et al*., 2005, 2009a, 2009b, 2009c; Jin *et al*., 2009) analyzed data from three independent experiments using a "mock theft" paradigm. Here, subjects were given a choice between taking a ring or a watch, with subjects stowing the selected item in a locker before scanning. During scanning, subjects were to deny possession of both items, rendering "lie" responses about the item taken and "truth" responses about the item not taken; baseline 'neutral' items elicited answers to general knowledge questions (*e.g*., 'Is it 2004?'; "Do you live in the United States?").

In all four Kozel *et al*. studies, the inferential logic was motivated by the analyses and findings in Kozel *et al*. (2005). In that study, data from an initial 30 subjects (Model-Building Group; MBG) served to identify brain regions that were more active in the lie versus truth contrast at the group-level.

[1] In addition to probing whether an individual responds truthfully vs. deceptively to forensically relevant questions, it appears that the instructed-lie ("stim test") conditions of the modified CIT paradigm of Hakun and colleagues (2008) are used by No Lie MRI, Inc. to validate that they can detect lie vs. truth telling in the individual. The findings of Hakun and colleagues raise doubts about the validity of this approach, because "deceptive behavior is not essential for the CIT-type response pattern…".

To develop and validate an individual-subject analysis, the researchers then computed, for each subject, the number of voxels (points in the brain) that were observed during the lie versus neutral contrast, and the number of voxels observed during the truth versus neutral contrast. Each voxel count was summed across three brain regions seen in the group-level analysis—anterior cingulate cortex, orbital frontal cortex, and inferior frontal cortex. If numerically more voxels were observed in the lie versus neutral contrast than in the truth versus neutral contrast for an individual, then it was concluded that fMRI detected the lie. If the number of voxels in the two contrasts were exactly the same, then the result was "inconclusive". If more voxels were observed in truth versus neutral than in lie versus neutral, then it was concluded that fMRI failed to detect the lie.

Having developed this approach using data from the MBG, Kozel and colleagues (2005) tested the approach on data from an independent group of subjects (Model-Testing Group). They observed 90% accuracy in determining whether a given subject had taken the ring or the watch (*i.e.*, the item taken was inferred by determining whether more voxels were active in the ring versus neutral contrast or in the watch versus neutral contrast). In subsequent papers, reanalyses of the data from Kozel *et al*. (2005) revealed that detection accuracy did not significantly improve (a) when adding galvanic skin response measures of autonomic arousal into the analysis (Kozel *et al*., 2009a), nor (b) when using a distributed pattern analysis approach (Jin *et al*., 2009).

Finally, two subsequent independent experiments using variants of the "mock theft" procedure revealed detection rates between 71% and 86% (Kozel *et al.*, 2009b, 2009c).

It might be tempting to conclude from these "mock theft" studies that fMRI-based methods can detect lies with modest to high accuracy (71%-90%). Unfortunately there is a fundamental lurking variable present in all of these experiments. Namely, the subject is likely to have richer memories for the object "stolen" than for the other object.

Specifically, in the "mock theft" procedure, the subject had an extended set of experiences with the "stolen" object (i.e., the subject selected the object, took the object from a drawer, and stowed the object in her/his locker, amongst the subject's personal belonging). By contrast, the subject had a more limited set of experiences with the other object (i.e., she/he viewed the non-selected object, but then left it in the drawer). Given these differences, it is likely that the subject has a richer set of memories for the experiences with the object chosen compared to the object not chosen. For this reason, the differences in the brain responses to

the two objects may not reflect deception per se, but rather could at least partially reflect memory effects.

Indeed, Gamer et al. (2009) recently demonstrated that memory processes, rather than deception, may account for group-level effects in some studies of deception. Moreover, as with the modified GKT/CIT paradigms (Hankun *et al.*, 2009), the selected object in the "mock theft" paradigm (*i.e.*, the object the subject is to 'lie' about) is likely to be more salient to the subject because it had been selected and acted upon, which raises the possibility that the observed brain effects also partially reflect differences in attentional orienting rather than differences due to deception.

Because it is unclear whether the published "mock theft" data reflect memory, attentional orienting, and/or deception, I conclude that *these studies do not provide unambiguous evidence bearing on whether fMRI-based methods can detect deception.* Future studies that eliminate the memory and attentional orienting confounds are needed before a conclusion can be drawn.

These "mock theft" studies also warrant an additional methodological comment. The conclusion about which object the subject took at the outset of the experiment, and thus which object the subject is lying about, rests on a difference between two voxel counts (i.e., the number of voxels showing a significant difference in the ring versus neutral contrast and in the watch versus neutral contrast), with relevant differences in the number of voxels in the two contrasts potentially being as small as a few voxels (or even just 1 voxel).

This creates the potential for an awkward (and prima facie invalid) situation. Specifically, imagine a subject who undergoes two fMRI scans (or two blocks within a single scan); during both scans she/he is asked to "lie," responding "no" to the object taken. Adopting the analysis logic implemented in the "mock theft" studies, if the investigator then counts the number of voxels that are more active during the lie versus neutral trials in the first scan and compares this count to the number of voxels that are more active during the lie versus neutral trials in the second scan, the investigator may be led astray.

Specifically, if the number of voxels is not exactly the same in the two contrasts (and there are many reasons why this is a low probability outcome when working with fMRI data), then the investigator will erroneously conclude that the subject was lying in one of the scans and telling the truth in the other, again, this will be the case even though the subject made the exact same response in the two scans. Similarly, this point applies to truthful answers to the other object—a comparison of the number of voxels across the two scans will lead to the

conclusion that the subject lied in one case and not in the other, even though the subject was truthful in both instances.

The central point here is that there are many reasons why the number of significantly activated voxels may differ between two comparisons, many of which have nothing to do with the psychology underlying the subject's behavior. It would appear problematic to rely on an analysis procedure that rests on an absolute comparison of the number of voxels between two contrasts in an individual subject, which assumes that a difference as small as a few voxels, or even a single voxel, is indicative of forensically important psychological and brain processes.

In addition to the fundamental methodological limitations noted above, it is important to emphasize that the preceding studies that aimed to assess fMRI-based lie detection at the individual-subject or individual-event level used tasks with low ecological validity, raising further uncertainty about their relevance for determining whether fMRI can detect lies in real-world situations.

In an effort to increase ecological validity, Kozel *et al*. (2009c) examined whether it is possible to use fMRI to identify which subjects had performed a more complex "mock sabotage crime" (damaging and stealing compact discs of incriminating video footage that the subject had watched) and which subjects had not performed this task. During scanning, subjects were instructed to deny performing these actions.

Using the same voxel-counting analysis procedure developed in the earlier ring-watch studies, Kozel and colleagues (2009c) correctly classified nine of nine subjects in the crime group as having performed the "mock sabotage crime" (100% sensitivity), but incorrectly classified ten of fifteen subjects in the no-crime group as having performed the crime (33% specificity). Given this low degree of specificity, these data indicate that it may be inappropriate to use this fMRI-based analysis approach for the detection of lies in the real world, as a low degree of specificity combined with a likely lower base rate of lying means that the number of false alarms (classifying someone as lying when they are being

truthful)may far exceed the number of hits (classifying someone as lying when they are being deceptive)[2].

Moreover, the same confounds of memory and event saliency that are present in the ring-watch "mock crime" paradigm are also present in this "mock sabotage crime" paradigm, which raises uncertainty whether the effects reflect deception, memory, and/or attentional processes.

For these reasons, the relevance of these data to answering whether fMRI-based methods can detect lies is unclear.

Conclusions

At present, the sensitivity and specificity of fMRI-based lie detection is unknown. Analysis of the published literature reveals no data that provides unambiguous evidence regarding the sensitivity and specificity of fMRI-based neuroscience methods in the detection of lies at the individual-subject or the individual-event levels.

While it is possible that fMRI methods will ultimately prove effective for lie detection, future studies are needed to eliminate fundamental confounds that exist in the published literature.

Additionally, other issues that are likely to prove important for forensic practice have received little to no attention in the literature (Greely & Illes, 2007). This includes

(a) whether the magnitude of the stakes of being caught lying matter,
(b) the effects of counter measures,
(c) how robust the methods are across subject populations (*e.g.*, older adults, individuals with psychiatric disorders, individuals taking medications, *etc.*),
(d) the effects of repeatedly probing the same "lie" or "truth" event,
(e) the effects of retention interval (time between the critical event and when the brain scans are conducted),

[2] Kozel et al. (2009c) conclude that their data indicate that their fMRI-based method "would be helpful to 'rule out' a potential suspect…" rather than to detect lying per se. Kozel et al.'s (2009c) observation that their method had 100% sensitivity and 33% specificity in this small sample size study (9 subjects in the crime group; 15 in the no crime group) appears to be the empirical basis for the conclusion that Cephos, Corp's method "is 100% accurate in determining truthfulness from a truthful person" (Kozel et al. 2009 is cited as the basis for this conclusion in an affidavit from Steven John Laken, Ph.D., President and C.E.O. of Cephos Corporation; February 4, 2010, United States of America v. Lorne Allan Semrau).

(f) the effects of instructed versus subject-chosen deception, and

(g) the effects of a lie's content (*i.e.*, what is being lied about).

Only future studies will tell whether fMRI-based neuroscience can identify lies.

References

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, *19*, 1557-1566.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W.,...Langleben, D.D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, *28*, 663-668.

Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2009). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social Cognitive and Affective Neuroscience*. Epub ahead of print (PMID: 19258375).

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex*, *13*, 830-836.

Greely, H.T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law and Medicine*, *33*, 377-431.

Hakun, J. G., Seelig, D., Ruparel, K., Loughead, J. W., Busch, E., Gur, R. C., & Langeleben, D. D. (2008). fMRI investigation of the cognitive structure of the concealed information test. *Neurocase*, *14*, 59-67.

Jin, B., Strasburger, A., Laken, S. J., Kozel, F. A., Johnson, K. A., George, M. S., & Lu, X. (2009). Feature selection for fMRI-based deception detection. *BMC Bioinformatics*, *10* Supplement 9, S15.

Kozel, F. A., Revell, L. J., Lorberbaum, J. P., Shastri, A., Elhai, J. D., Horner, M. D.,...George, M. S. (2004a). A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men. *Journal of Neuropsychiatry and Clinical Neurosciences*, *16*, 295-305.

Kozel, F. A., Padgett, T. M., & George, M. S. (2004b). A replication study of the neural correlates of deception. *Behavioral Neuroscience*, *118*, 852-856.

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, *58*, 605-613.

Kozel, F. A., Johnson, K. A., Laken, S. J., Grenesko, E. L., Smith, J. A., Walker, J., & George, M. S. (2009a). Can simultaneously acquired electrodermal activity improve accuracy of fMRI detection of deception? *Social Neuroscience*, *4*, 510-517.

Kozel, F. A., Laken, S. J., Johnson, K. A., Boren, B., Mapes, K. S., Morgan, P. S., & George, M. S.(2009b). Replication of functional MRI detection of deception. *Open Forensic Science Journal*, *2*, 6-11.

Kozel, F. A., Johnson, K. A., Grenesko, E. L., Laken, S. J., Kose, S., Lu, X.,...George, M. S. (2009c). Functional MRI detection of deception after committing a mock sabotage crime. *Journal of Forensic Science*, *54*, 220-231.

Langeleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*, 262-272.

Monteleone, G. T., Phan, K. L., Nusbaum, H. C., Fitzgerald, D., & Irick, J. S. (2009). Detection of deception using fMRI: Better than chance, but well below perfection. *Social Neuroscience*, *4*, 528-538.

What is neurogenetics?

Louis J. Ptáček, M.D.

Neurogenetics is the field within genetics which examines the impact of genes on the structure and function of the brain and nervous system. Information gained in this rapidly developing field assists scientists and medical doctors in identifying factors related to diseases and pathologies located within the brain and nervous system. Neurogenetics also identifies genes associated with positive traits, such as musical talent, and negative traits, such as violence and aggression.

Genetic evidence may be introduced to a court for a variety of reasons. Defendants may support an insanity defense with this information. Prosecutors may assert that certain individuals present a more grave danger to society based on brain conditions demonstrated through genetic information. Such evidence may relate to genetic evidence about a party who alleges he or she has a certain inherited disease or it may be based on the statistical correlation of a certain genetic trait with a disease or disorder.

Types of Neurogenetic Information

Understanding the differences among the many kinds of genetic data is critical to an appreciation for what such data demonstrate. Some disorders may be followed as they are passed down through families. They are referred to as disorders that "segregate." The simplest forms of these disorders are called "Mendelian traits," after the nineteenth century Augustinian monk Gregor Mendel.

The three main patterns of segregation are termed autosomal dominant, autosomal recessive, and X-linked disorders. In autosomal dominant diseases, a single copy of a mutation inherited from either parent is sufficient to cause a disease. A single mutant form of the gene is dominant over the normal gene on the other chromosome. Typically, any child who has a parent with an autosomal dominant disease has a fifty percent chance of inheriting that disease.

In contrast, autosomal recessive diseases tend to involve the loss of one functional copy of a gene. Because the gene on the other chromosome is functional, those carrying the recessive gene do not manifest a disease. However, if two individuals with an autosomal recessive

gene have a child together, there is a 25% chance that the child will inherit two nonfunctional copies of the gene and exhibit the disease.

Autosomal dominant	Single copy of a mutation from one parent causes disease (except in rare cases called non-penetrance)
Autosomal recessive	Loss of one functional copy of gene inherited from one parent not sufficient for disease (but two such individuals with recessive genes have 25% chance of having child with disease)

Even in autosomal dominant diseases, it sometimes occurs that a person carrying the mutation does not manifest the disease. This phenomenon is called non-penetrance. Some portion of the population may, in fact, have the genetic variation but show no outward signs of disease or disorder. There are a number of possible explanations for this. Certain genetic diseases and disorders become apparent later in life. Alternatively, other genetic factors within the individual could be offsetting the mutation or certain environmental exposure helped protect against the manifestation of the disease.

Of course, there is a spectrum of penetrance. When geneticists describe a disease as highly penetrant, they mean to indicate that most or all individuals with the mutation show the disease. Certain autosomal diseases, which one would expect to be manifest, present varying degrees of penetrance. In certain circumstances, for example, 5% of all individuals with the autosomal dominant gene may not express the disease normally associated with the mutation. At the other end of the spectrum, a combination of multiple variants is required to add up to a sufficient genetic "burden of disease." Such diseases are polygenetic traits.

Expressivity is another important term and refers to the severity of disease in individuals with a genetic disease. Penetrance can be high with variable expressivity. This could lead,

for example, to a family where most or all gene carriers have aspects of the disease but where some are very mildly affected and others are severely affected. The degree to which individuals who express a disease are affected is called "variable expressivity."

Great differences exist between the various genetic mechanisms discussed in this section. Additionally, it is impossible to draw a discrete line between Mendelian disorders with incomplete penetrance and those disorders with polygenetic traits requiring a mixture of genetic factors to combine in order to manifest.

Rounding out this description of genetic disorders are those labeled complex genetic diseases. These are disorders in which particular genetic variants do not cause disease but instead are associated with an increased risk of having that disease. There is a fundamental difference between gene variants that cause human diseases with high likelihood and those that increase risk for human diseases.

Statistical Association

Neurogenetics looks not just at genetic mutation in individuals or families (where mutations could be said to "cause" disease) but also at the correlations between certain genetic markers and disease in larger populations. Geneticists look at the statistical association of genetic markers with human diseases. The most well known example of such an association is that the Apo E allele is "associated" with Alzheimer's disease. Such associations are found by looking at large patient populations and comparing the patterns of genetic variation with a large group of matched controls.

While particular genetic variants may increase the risk of a particular disease, carrying that variant does not guarantee that one will develop the disease. Furthermore, not carrying the variant does not guarantee that an individual will be free of the disease.

Statistical association research could begin, for example, with a group of 1,000 unrelated patients with Alzheimer's disease from whom DNA samples have been obtained. A control population would be composed of the same number of ethnically similar individuals who are gender and age matched to the patients in the data set with Alzheimer's disease.

If a genetic variant is found in 30% of Alzheimer patients but only 10% of the control subjects, researchers can conclude that one carries that genetic variant. Such variants are not "causative." Rather, they are associated with an increased risk of developing Alzheimer's disease. However, 70% of the Alzheimer patients in this example do not have the genetic

marker that is associated with disease. Furthermore, 10% of the control populations who don't have Alzheimer's disease still carry the risk allele. In this situation, one cannot be certain that an individual who is carrying a marker of increased risk will manifest the disease or that someone who is not carrying that marker will be free from disorder.

While the genetic variant changes the risk that a particular person may develop a disease, it is frequently the case that the genetic variant that is serving as the marker is not itself contributing to the increased risk. A genetic marker from a human chromosome map might be close to the gene that contributes to increased risk but itself is not contributing to that increased risk.

Examining Chromosomes

One of the earliest insights into the basis of genetic disorders was observed when it became possible to look at smears of chromosomes under the microscope. A chromosome is a threadlike strand of DNA in the cell nucleus that carries the genes in a linear order. Analysis of karyotypes (a term for the number, form, and size of chromosomes) allowed visualization of large deletions of parts of chromosomes, insertions of large pieces of DNA into a chromosome, translocation of a part of one chromosome onto another, or the presence of less than or more than two chromosomes for any pair.

The most common example of the latter is trisomy21 where three copies of chromosome 21 are observed to be present in individuals with Down Syndrome. Individuals with classic features of Down Syndrome carry three copies of chromosome 21 or two copies of chromosome 21 and a partial copy of a third chromosome 21. Using this information, it is possible in this case to pre-symptomatically diagnose a fetus by examining karyotypes on cells from the fetus. It is also possible to perform such tests postnatally.

While Down Syndrome is the most common of this group of disorders, there are numerous other conditions that are known to be caused by the loss of one copy or the gain of one or more additional copies of individual chromosomes (or parts of chromosomes). In such cases, it is possible to say that the abnormal number of chromosomes is causative for the disorder.

Genes or Nurture?

People have argued for a very long time whether certain traits arise from genetic factors or environmental factors. We can end this debate because we now know that all disorders (and traits) are affected by both genes and the environment. For example, Huntington's disease

(a neurodegenerative genetic disorder) is a highly penetrant autosomal dominant Mendelian trait. As discussed above, this means that it is passed from only one parent with the genetic variation and that everyone who has that variation expresses Huntington's disease.

Even in a case of what appears to be overwhelmingly strong genetic causation, two individuals with the identical mutation may manifest the disease at different points in their lives. This may be caused by other genetic variables and to environmental exposure.

For example, in the case of two brothers with the mutation, one may lead a life of physical activity, eat healthy foods, and refrain from smoking and alcohol consumption. The other brother may be sedentary, unhealthy, and alcoholic. The variation in manifestation of the disease would not surprise us given these facts.

While Huntington's disease is identified as a disease almost solely dependent on a genetic mutation, HIV infection is often portrayed as a disease based on environmental factors. Individuals who subject themselves to certain environmental factors (for example, unsafe sex and intravenous drug use) have a higher risk of developing the disorder. Even here, though, genetics matter.

Among those who are infected with HIV, there is a group of individuals who are very long survivors even in the absence of treatment. Some of these individuals have been shown to have genetic variants in chemokine receptors (CCR5 and CCR2). The interaction of genetic and environmental factors, exemplified in these two cases, seriously complicates the interpretation of genetic associations discussed above.

Distinguishing the Data and Applying Neurogentics to Law

As mentioned above, attorneys may currently (or will soon) introduce neurogenetic information to "explain" certain behaviors. Apart from the scientific and legal issues associated with the question of whether identifying the causes of certain behaviors is relevant to the law, judges should consider two types of neurogenetic information.

It is crucial for legal actors to distinguish between the Mendelian genetic variations that "segregate" down through families and the association of certain genetic variations with various disorders. The Mendelian genetic variations are often said to "cause" a disorder. Statistical association, in contrast, may provide only a correlation between a certain

variation and a group with a disorder. The genetic variation may have nothing to do with the disease itself (it may simply be a marker related to, but not causing, the disorder).

Furthermore, statistical association cannot be used to show that a certain person has developed or will develop a disorder. Nor can it prove that those without the variation will not develop that disorder.

References

Watson, J.D., Myers, R.M., Witkowski, J.A., Caudy, A.A., (2006). *Recombinant DNA: Genes and Genomics: A Short Course* (3rd Edition), Chapters 1 and 2. W. H. & Company, 2006.

Rosenberg, R. N., DiMauro, S., Paulson, H. L., Ptáček, L., and Nestler, E. J. (2007). *The Molecular and Genetic Basis of Neurologic and Psychiatric Disease* (4th Edition). Wolter's Kluwer/Lippincott Williams & Wilkins.

Jorde, L. B., Carey, J. C., Bamshad, M. J., White, R. L. (2000). *Medical Genetics* (2nd Edition). Mosby-Year Book.

Can neuroscience identify pain?

Howard Fields, M.D., Ph.D.

We are being asked whether it is currently technically feasible to peer into the subjective experience of an individual claiming to experience pain. The straightforward answer is yes and no. Currently we do not have any devices that would allow us to monitor the conscious thoughts of another individual; we have direct access only to our own subjective experience.

On the other hand, it is reasonable to ask whether we know enough about the neural processing of pain to predict that an individual is experiencing pain when a specific pattern of neural activity is observed using objective measures. The weight of evidence supports the notion that it is, in theory, possible to do this. However, to my knowledge, no direct experimental test of this approach has yet been carried out.

Using a variety of experimental approaches, we have achieved a broad and deep understanding of pain processing (McMahon & Koltzenburg, 2006). Work in animals has identified the nerve cells that innervate pain sensitive tissues. We have good insight into the molecular mechanisms by which these nerve cells sense tissue damage and how the pain message is encoded in the firing pattern of the nerve fibers that connect the sensitive tissues to the spinal cord. We have also mapped the responses of central nervous system neurons in a variety of species, including members of our own biological order, the primates. The pain pathway from spinal cord to the brain is well understood in animals.

Furthermore, although technical and ethical barriers prevent us from studying the human central nervous system at the single cell level except under the most restricted circumstances, what information we do have indicates that humans and their close relatives in the animal kingdom process pain in the same way.

Parallel studies in humans and primates indicate that pain is encoded in the frequency of firing of pain sensitive nerves in the periphery, in the spinal cord, and in the thalamus region of the brain (the first stop for the major pathway that conducts the pain message from the spinal cord, the spinothalamic tract). In human studies, electrical stimulation in the appropriate thalamic regions cause pain, and lesions in this region produce a permanent loss of normal pain sensation.

The biggest breakthrough in relating human brain activity to pain has depended on functional magnetic resonance imaging (fMRI) (*see* Raichle, above*)*. The fMRI signal has fairly good spatial resolution, but it is slower than the actual neural activity that induces the blood flow changes. It is important to point out that the fMRI BOLD signal measures changes in blood flow and so the signal during the painful stimulus is compared to a baseline period when no pain stimulus is applied. The form of the data is the anatomical pattern of BOLD signal contrasts between a painful stimulus and no stimulus (or between a low intensity innocuous stimulus and a painful stimulus).

There is broad agreement that application of a noxious stimulus induces BOLD signal changes in a specific anatomical pattern in the human brain, and that the magnitude of these changes shows a significant correlation with an individual's subjective rating of pain intensity. Furthermore, invasive animal studies are consistent with the idea that individual nerve cells in the regions showing the BOLD signal in people encode pain intensity (*i.e.*, their discharge rate correlates with the stimulus intensity in a range of intensities that people report as painful).

Thus, in theory, if we know exactly when to measure the BOLD signal in the brain (*e.g.*, a specific time following a stimulus, like a pinch or a needle prick) we should be able to use that signal to reliably predict whether an individual was going to report the stimulus as painful. Unfortunately, technically, we aren't quite there yet.

The major problem is that the BOLD signal following a single painful stimulus is very small, so that stimuli have to be applied multiple times to get a signal that can be clearly distinguished from noise (signal averaging). Even with signal averaging, most fMRI studies have to pool data from multiple subjects to get a signal that can be reliably correlated with pain perception. I have no doubt that this technical problem will be solved, but it will take a few more years.

A second more practical problem relates to how the imaging would be used by lawyers and judges. It is unlikely that they would be interested in the pattern of activity produced by a precise series of calibrated stimuli carried out with informed consent in a laboratory setting. The goal of those laboratory studies is to determine the brain mechanisms underlying pain perception.

Most likely, the legal profession would be more interested in an objective measure of a pain that occurred following an injury or that results from a disease process. There the goal

would be to determine degree of harm to the plaintiff and the liability of a defendant. Obviously, if a person's compensation to make them whole under the law depends upon their report of pain, they would have an incentive to report a higher level of pain, i.e. they would have a conflict between self-interest and accuracy.

It would be nice to have an objective measure that is independent of the subject's bias; something equivalent to an x-ray or a blood test that can objectively reveal organ disease.

Is such a measure possible? The difficulty is that in order to relate the BOLD signal to pain, you would need to know what the person is feeling at the time of the scan. This is challenging when there is no experimenter applying a painful stimulus at a known time. To determine the pattern of activity, the fMRI BOLD signal during pain has to be compared with a baseline condition when there is no pain (or a known but different level of pain). However, to do this comparison you need the subjects' report of their own pain levels. Is this possible?

It turns out that this is, hypothetically, possible. Because we know where to look in the brain for activity induced by painful stimuli, we can question the subject about their current ongoing pain level. It turns out that in most painful conditions patients reported pain levels that vary in intensity over time.

With this in mind, Apkarian and his colleagues at Northwestern University took an ingenious approach (Baliki *et al.*, 2006). They placed subjects in the fMRI magnet and had them rate their spontaneous pain level continuously using an electronic indicator. While the subjects were providing their pain ratings, their brains were intermittently scanned. Apkarian and his colleagues were able to show that when the subjects reported their pain as rising in intensity, the pattern of fMRI BOLD signal was very similar to that seen in studies of normal subjects during application of a controlled painful stimulus.

In theory, this approach could be used to confirm that an individual who was reporting a given level of pain was experiencing it due to activity in the traditional pain pathways. This would provide objective support for the subject's testimony. This method could be made more quantitative by using it in combination with measurements of the fMRI BOLD during the application of a set of calibrated noxious stimuli of varying intensity to the same individual who is scanned at a different time, while reporting the ongoing spontaneous pain related to the legal issue.

Applying painful stimuli of increasing intensity to an individual should produce a BOLD signal of correspondingly increasing intensity. In this way, for a given individual, it might be possible to calibrate the subject's own pain related BOLD signal in a specific set of brain areas. The BOLD signal induced by the known painful stimulus could then be compared to the signal seen during the subject's report of spontaneous pain. Thus the BOLD signal could be used not only to affirm that an individual is indeed feeling pain, but could, additionally, provide an objective correlate of its perceived intensity.

One recent study has provided an unexpected but very robust measure that appears to separate normal subjects and those reporting moderate to severe ongoing pain. Baliki and colleagues, again at Northwestern, compared fMRI BOLD imaging in normal (pain free) subjects and those with chronic low back pain (Baliki *et al*, 2010). An experimental heat pain stimulus of varying intensity was applied to the skin over the back to both groups of subjects. The BOLD signal was identical in the traditional pain responsive areas in both groups of subjects, however, when the applied experimental painful stimulus was turned off, there was one brain region that showed an opposite response in normals and in subjects with ongoing low back pain; the ventral striatum. Activity in this region has been implicated in both animals and humans as signaling pleasure or reward.

In normal subjects, turning off the painful stimulus was experienced as rewarding and was associated with a positive BOLD signal. In contrast, in the pain patients, the acute heat pain stimulus had relieved their ongoing back pain and when it was turned off, their back pain returned, so they felt, overall worse, and the BOLD signal in this area was reversed to negative. Because the BOLD responses to pain offset in normal and low back pain patients were non-overlapping and of opposite sign, this might be the most technically accessible fMRI signal that an individual has ongoing pain.

So to the question, "Can neuroscience identify pain?" the answer is in theory yes. We can show its neural correlates in conscious human subjects but only using signal averaging and groups of subjects responding to applied stimuli. The signal is there, but it is small, and currently has not been shown to be useful in an individual at a single moment in time. It seems feasible that with longer periods of measurement or with improvements in signal to noise these problems can be overcome. It is important to point out that while the methods are robust, they are indirect. Pain depends upon activity in a specific set of neurons, their activation is sufficient to produce it, but in the end, the experience itself is subjective and therefore our methods to measure it are, of necessity, indirect.

References

Baliki, M. N., Chialvo, D. R., Geha, P. Y., Levy, R. M., Harden, R. N., Parrish, T. B., & Apkarian, A. V. (2006). Chronic pain and the emotional brain: Specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *Journal of Neuroscience*, *26*, 12165-12173.

Baliki, M. N., Geha, P. Y., Fields, H. L, & Apkarian, A.V. (2010). Predicting value of pain and analgesia: Nucleus accumbens response to noxious stimuli changes in the presence of chronic pain. *Neuron*, 66, 149-160.

McMahon, S. B., & Koltzenburg, M. (2006). *Wall and Melzack's Textbook of Pain* (5th Edition). London: Elsevier Churchill-Livingstone.

Does neuroscience give us new insights into criminal responsibility?

Helen Mayberg, M.D.

Advances in neuroscience have revolutionized the exploration of the micro- and macro-organization of the central nervous system revealing an increasing complex anatomical, physiological, biochemical and molecular organizational structure.

Progress in neuroimaging has facilitated new research investigations into normal human brain functioning and have provided important new insights into the mechanisms of many neurological and psychiatric disorders with associated implications for diagnosis, treatment and risk assessment.

As outlined by previous contributors, state-of-the-art structural magnetic resonance imaging (sMRI) allows precise measurements of focal brain lesions as well as subtle changes in brain shape or volume over time. This provides methods for direct correlations with specific clinical abnormalities, building on classical pathological lesion-deficit studies and extending to studies of normal variation dictated by gender and genetics.

Similarly, the newest functional imaging methods including positron emission tomography (PET), electroencephalography and event related potentials (EEG/ERP), and functional MRI (fMRI) not only provide strategies to examine regional abnormalities in well-characterized brain diseases, but have expanded the range of testable hypotheses about normal sensory, motor, affective and cognitive processes.

Experiments examining reward valuation, risk assessment, moral judgment, stress effects on decision making, and responses to peer pressure, among many inventive paradigms have identified various findings in healthy subjects. This allows us to establish potential criteria for defining "deviations" from "normal" using increasingly complex behavioral criteria. Classifier analyses applied to performance of explicit tasks can even make accurate predictions of a subject "decisions" on a trial by trial basis during such experiments.

At issue here is whether such research strategies and discoveries can be generalized for use in the evaluation of an individual in the setting of a singular set of circumstances associated with the purported commission of a specific criminal act. More specifically, can

neuroscience investigations provide insights into a given defendant's intent and motivation at the time of a given crime?

Generally, any data regarding brain structure or function is unlikely sufficiently contemporaneous to the time of the crime to be meaningful. At best, and presuming that the presented data meets evidentiary standards of admissibility, these data can only provide general inferences about the relationship of any identified "abnormalities" in brain structure or function to a person's behavior.

Such findings cannot provide conclusive information about specific behaviors which happened in the past, nor predict likely behaviors in the future. For instance, research studies of clinically well-characterized psychopaths have reported structural and functional abnormalities (reviewed in Yang & Raine, 2009), while performing explicit behavioral tasks of relevance to the investigation of psychopathy.

While these type of functional findings may inform us about potential brain mechanisms mediating specific cross-sectional behavioral differences between psychopaths and healthy volunteers, such findings cannot inform on the likely presence or absence of that behavioral aberration (or by inference, any other behavior aberration) in an individual psychopath (or an individual criminal defendant without that diagnosis). This is especially true regarding likely behavior at a specific past point in time under the unique set of circumstances surrounding their specific criminal act.

Therefore, these sorts of research findings, no matter how compelling, contribute little to determining *mens rea* in a given defendant.

That is not to dismiss ongoing studies investigating brain abnormalities in subjects with well-characterized behavior disorders, including antisocial personality or psychopathy. Such investigations have been the logical next step from the many compelling case reports and population studies demonstrating consistent relationships between certain brain lesions (orbital frontal trauma, surgical resections, invasive tumors), post-infectious processes (limbic encephalitis), and neurodegenerative disorders (fronto-temporal dementia), on one hand, and acquired behavioral disturbances affecting emotion regulation, decision making and impulsive control, on the other.

These findings are demonstrated in the work of Antonio Damasio, Daniel Tranel, James Grafman, and others. Other researchers, such as Kent Kiehl, Adrian Raine, and Robert Hare,

note that dramatic and pervasive behavioral changes following focal lesions invite a search for similar, albeit more subtle, anatomical and functional abnormalities in criminal defendants without clear cut injuries accused of committing violent crimes, and in individuals classified as antisocial or psychopathic.

Findings implicate abnormalities in various parts of frontal cortex, seemingly consistent with focal lesions studies. The presence of such lesions, however, does not reliably predict the presence of such behavioral abnormalities in any individual. Further, the reliability of these reported scan patterns for diagnosing psychopathy, for example, is not yet established and error rates (false positive and false negative) are unknown. Interestingly, most patients with these type of lesions do not display antisocial or criminal behavior and not all criminals show such brain abnormalities.

At present, all available functional neuroimaging procedures index neural activity during acquisition of the scan. Therefore, findings give information within that experimental context. While a scan pattern might reliably correlate with the presence of a pathological condition, it cannot predict the status of that condition, such as the presence or severity of symptoms, or likely progression or prognosis of the condition. By way of example, metabolic abnormalities can be identified using FDG (fluorodeoxyglucose) PET scanning in individuals with genetic risk for Alzheimer's disease but who show no symptoms of any type, as shown in the work of Eric Reiman. What then could one predict or infer about said individual's mental state even at the time of the scan, much less during the past?

Similarly, one cannot assume stability of a given functional scan finding, particularly if that type of injury or condition has not been shown to leave static scan abnormalities of a known pattern (frontal hypometabolism in depressed patients resolves with remission of clinical symptoms, for example).

At the point that a scan "finding" is identified as purportedly linked to a given condition, other concurrent, preexisting, or newly discovered neurological or psychiatric conditions or symptoms must be considered. Moreover, use of psychoactive medications like sleep, anti-epileptic, antidepressant, and anti-anxiety medications, as well as the patient's behavioral state, mood and motivation at the time of scanning (anxious, sad, sleepy, distracted, uncooperative), must also be considered as potential contributors to any observed deviant scan pattern.

The independent impact of prosecution on a defendant's activity patterns during a scanning session, as compared to paid healthy volunteers, may also be a confound. Distinguishing and disassociating scan patterns and abnormalities due to such factors from those due to the condition in question is not possible, even when there is available data regarding these factors on the day of the imaging study.

Generally, when considering the diagnostic utility of any test (be it imaging or any other method), one must first address if the behavioral state under investigation is static (developmental anomaly, old head injury), episodic (bipolar manic versus depressive versus euthymic state), or progressive (Alzheimer's disease, fronto-temporal dementia).

From the perspective of disease diagnosis, scan abnormalities (*i.e.*, disease patterns) must first be identified and statistically confirmed in groups of subjects with the identified condition verified using independent clinical and pathological criteria. A scientifically valid correlation must then be shown between the scan pattern and the independent clinical/pathological criteria before the significance of specific abnormalities can be attributed to a specific disease or clinical syndrome or condition. Lastly, group patterns must be shown to be reliably detectable in individual subjects, including a determination of sensitivity and specificity (error rate, false positives and false negatives).

Sensitivity and specificity of a given pattern for a given diagnosis (or condition or behavior state) will vary by procedure and diagnosis, but can be quantified. However, what is considered an acceptable error rate is more ambiguous, being dependent on consideration of the consequences of a false positive versus a false negative conclusion. That said, once such reliable findings or patterns have been established through replicated experimental studies by multiple investigators, scans performed on any individual can be theoretically interpreted with the following criteria:

- Is the pattern of brain activity in Subject X significantly different from the pattern seen in persons who by history and clinical exam are deemed to be "normal?"
- Does the pattern seen in Subject X significantly deviate from "normal" in a manner consistent with (or matching) one previously recognized for a specific pathological condition?

Even if such criteria can be met, further inferences as to the likely behavioral state present at the time of the crime are impossible to infer from any scan pattern.

In conclusion, neuroscience studies are making systematic advances in understanding increasingly complex aspects of human behavior, including probabilistic likelihood of particular responses during real-world experimental paradigms.

Despite these advances, at our current state of knowledge, it is beyond the data generated from any currently published scanning protocol to make predictions about the rational capacity (or lack thereof) of a criminal defendant, or to make inferences as to that defendant's intent at a specific moment in time before or during a specific criminal act. Time will tell if paradigms can be designed that meet the necessary criteria to make such inferences.

References

Yang, Y., & Raine, A. (2009). Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: A meta-analysis. *Psychiatry Research*, *174*, 81-88.

Müller, J. L., Sommer, M., Döhnel, K., Weber, T., Schmidt-Wilcke, T., & Hajak, G. (2008). Disturbed prefrontal and temporal brain function during emotion and cognition interaction in criminal psychopathy. *Behavioral Sciences and the Law*, *26*, 131-150.

Howard, R., & McCullagh, P. (2007). Neuroaffective processing in criminal psychopaths: brain event-related potentials reveal task-specific anomalies. *Journal of Personality Disorders*, *21*, 322-339.

Batts, S. (2009). Brain lesions and their implications in criminal responsibility. *Behavioral Sciences and the Law, 27*, 261-272.

Garland, B., & Glimcher, P. W. (2006). Cognitive neuroscience and the law. *Current Opinion in Neurobiology*, *16*, 130-134.

Morse, S. (2005). Brain overclaim syndrome and criminal responsibility: A diagnostic note. *Ohio State Journal of Criminal Law*, *3*, 397-412.

Mayberg, H. S. (1992). Functional brain scans as evidence in criminal court: An argument for caution. *Journal of Nuclear Medicine*, *33*, 18N-19N, 25N.

Does neuroscience give us new insights into drug addiction?

Floyd E. Bloom, M.D.

Becoming addicted means a life of compulsive drug seeking and use, despite the severely negative consequences of such self-medication. Research using experimental animals, from rodents to non-human primates, has shown that normal animals, never previously exposed to addictive drugs, will readily self-administer every drug that human beings abuse. When these animals are given access to these drugs of abuse (alcohol, nicotine, heroin, cocaine, and amphetamines), they will continue to self-administer these drugs compulsively.

Neuroscience research has shed new light on the biological underpinnings of drug addiction, allowing researchers to devise new interventions and develop new treatments for addictive disorders.

What is addiction?

For more than forty years, it has been known that experimental animals will do work (press levers or poke their noses in holes) in order to activate electrodes implanted in specific regions of their brains. The spots that elicit such self-rewarding behavior served to help us identify an internal reward system in the brain. This series of interconnected brain regions normally functions to reinforce life-sustaining drives, such as thirst, appetite, and reproductive behaviors.

While the anatomical circuitry associated with reward has been known for some time, modern neuroscience research has identified the chemical signaling systems used by these pathways, and specified how the drugs of abuse act to activate the brain's reward system and mislead the brain into identifying the use of the drugs as a functional reward.

This reward pathway consists of neural connections between the ventral tegmental area and the nucleus accumbens, and contains the monoamine neurotransmitters associated with mood. Addictive drugs act in the brain by increasing the interneuronal signals of dopamine, norepinephrine, or the naturally occurring endogenous opioid in the reward pathway. This increased cellular reward signaling produces a reinforcing effect on the addictive behavior.

Continuous use of addictive drugs causes the brain to take adaptive steps to overcome the effects of the drugs. The brain performs these adaptations by making active those circuits

whose effects oppose the sedating, stimulating, or mood altering effects of the abused substance. When the dependent person stops self-administering their abused drug, the overactivity in these adaptive opposing circuits continues, producing the signs and symptoms of withdrawal, and inducing the drug user to reinitiate drug use in order to suppress the withdrawal symptoms.

After understanding the biological underpinnings of addiction, researchers were able to use these animal models to devise treatments for nicotine dependence, opiate addiction and alcohol dependence. Nicotine replacement products such as chewing gums, and skin patches, release nicotine into the blood stream at levels that prevent the appearance of the withdrawal symptoms. This respite from withdrawal that would otherwise coerce further tobacco use, allows the persistent patient the time required to reduce the counter-drug adaptive processes, and restore a healthful condition.

For opiate and alcohol dependence, the appropriate brain receptors for the endogenous opiate transmitters can be occupied by the drug naloxone. Naloxone acts to block the drug effects of opioids, such as heroin, by blocking opiate receptors in the brain. This treatment has been used acutely in cases of respiratory depression in infants born to addicted mothers and in cases of morphine or heroin overdose, to reverse some of the harmful effects of heroin.

Long acting forms of naloxone can provide opiate antagonism for weeks, increasing the time of respite from withdrawal without requiring the compliance of the addict to take the drug. These advances in our understanding of addiction are leading to the development of fundamentally new treatments for addictive disorders that are already under clinical testing.

Why do some become addicted and some do not?

Drug addiction is a chronically relapsing disorder characterized by a compulsion to seek and take a drug, loss of control in limiting intake, and emergence of a negative emotional state (for example, dysphoria, anxiety and irritability) when access to the drug is prevented. An important goal of current neurobiological research is to understand the molecular, neuropharmacological, and neurocircuitry changes that mediate the transition from occasional, controlled drug use to the loss of behavioral control over drug seeking and drug taking that defines chronic addiction. For much of the twentieth century, drug addiction was regarded as a personality issue, as a habit that the addict could break if they had sufficient will power to do so.

However, beginning in the 1970s, solid research in humans and in animal models of addiction indicated that a vulnerability to becoming addicted was biologically based and inheritable. Lines of animals who were vulnerable or resistant to drug self administration were created by inbreeding, while human research indicated that children of alcoholics adopted away from their dependent parents expressed the same higher levels of drug dependence as those raised by the addicted parents.

The modern molecular research that has identified the neurotransmitters systems that underlie the specific addictive effects of opiates and alcohol has also shown that those individuals whose opiate receptors are somewhat less sensitive to opiates, especially among Caucasians, are more vulnerable to opiates and heroin addiction and more readily treatable by the opiate antagonist drug naltrexone. As this research reveals more neurotransmitter involvement in the addictive process, it is likely that additional interventions will become manifest. These biological vulnerabilities do not exonerate the person for responsibility for their addictive state since it is their choice to use the drugs, once or multiple times.

Should we punish addicts?

During the 1970s, and indeed even occasionally today, drug addiction was considered by law enforcement officers and the criminal justice system to be instant and permanent, inducing a craving so powerful that no conscious effort could overcome it. For those addicts in withdrawal, overtly criminal behavior to acquire drugs was considered justifiable.

However, research with large samples of soldiers, based on testing and interviews one and three years after their military service, provides astounding results. Initial interviews supported by urine testing indicated that nearly 80% had used marijuana, half of all enlisted men had tried morphine or opium, and that nearly 20% were symptomatic enough to have been called dependent while in the service. One year later only 5% of those who were addicted to opiates in the war zone were addicted in the United States. Of those not addicted, virtually none had received any treatment.

Lee Robbins of Washington University in St. Louis, the lead epidemiologist of those studies, concluded that the availability of cheap drugs accounted for the high rates of drug use in wartime. Clearly, the common view of the addict—once addicted, addicted for life—was erroneous. Addiction was not a lifelong dependency; it could be interrupted by a change in environment even without treatment. Perhaps, with the right agent, treatment was possible.

However, for the veterans who exhibited deviant social behavior before serving in Vietnam, the rates of re-addiction and treatment failure were as high as in the civilian and federal prison populations. In the case of alcohol dependence, the lifetime prevalence approaches 20% in the general population. To imprison an individual and provide neither treatment nor the prevention of access to the drugs to which they were dependent before imprisonment sacrifices all the knowledge that an addict can be treated and leave prison free of drugs.

Effective treatments and interventions are available, although psychostimulant dependence remains an area of intense research development.

Death from opiate overdose is a major source of mortality following release from incarceration. A prerelease program of education and the provision of an opiate antagonist, such as naltrexone, to the released prisoners helps reduce mortality.

Moreover several treatments for opiate and alcohol dependence have become available to physicians treating addicts. One, Vivitrol™, is available in a long lasting form requiring one injection a month thereby eliminating concerns for compliance. This appears to be a good alternative to an oral medication that needs to be taken one or more times daily, such as with alternate treatments like buprenorphine or acamprosate. It is even possible to predict effectiveness of this treatment option with genetic testing for alternative forms of the opiate receptor where both alcohol and opiate drugs act. In either alternative, treatment with behavioral therapy has been shown to be an important complement to medication treatment for addiction. Treatment for marijuana dependence through various medications is almost as effective as pharmaceutical treatment for dependence on opiates and alcohol.

Basic neuroscience research strongly supports the position that an untreated addict released into the social environment in which their drug use was previously undertaken will almost certainly result in a return to drug use and the accompanying criminal activity undertaken to support it. Prisoners should be treated for their addictions, given a respite from the ability to obtain the drug of choice (or any drugs), and returned to society in an environment sufficiently different from the one in which they were dependent to help break the addiction.

References

Taxman, F. S., Henderson, C. E., & Belenko, S. (2009). Organizational context, systems change, and adopting treatment delivery systems in the criminal justice system. *Drug and Alcohol Dependence*, *103S*, S1–S6.

Henderson, C. E., & Taxman, F. S. (2009). Competing values among criminal justice administrators: The importance of substance abuse treatment. *Drug and Alcohol Dependence*, *103S*, S7-16.

McCarty, D., & Chandler, R.K. (2009). Understanding the importance of organizational and system variables on addiction treatment services within criminal justice settings. *Drug and Alcohol Dependence*, *103S*, S91-93.

Can neuroscience identify psychopaths?

Kent Kiehl, Ph.D.

Psychopaths constitute less than 1% of the general population, but they commit over 30% of violent crime in our society and create an estimated $250-400 billion dollars in cost to society each year (Reznek 1997). Hence, they pose a central problem for the criminal justice system.

The modern concept of psychopathy was formulated by French psychiatrist Pinel (1792), refined through 40 years of clinical case studies by American psychiatrist Hervey Cleckley (Cleckley, 1941, 1976), and then operationally defined by Canadian Robert Hare and colleagues in the Hare Psychopathy Checklist (Hare, 1980) and its successor, the Hare Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003; see Table 1). There is a substantial literature attesting to the reliability and validity of the PCL-R as a measure of psychopathy in offenders and forensic patients (see Hare, 2003 for review). The PCL-R is now the most widely accepted diagnostic instrument for psychopathy in general, and in legal contexts in particular. It is the gold standard for this diagnosis.

The PCL-R is administered by trained personnel. They begin with a comprehensive file review on the client, including information on education, employment, relationships (family and romantic), drug/alcohol use, and lifetime criminal behaviors derived from both previous interviews with the client and collateral reports from family, friends, police, victims, etc. The file review is typically followed by a 2-4 hour interview with the client covering all areas of the subject's life.

After completing the file review and interview, the subject is rated on each of the 20 items of psychopathy (Table 1). Each of the 20 items is scored according to explicit criteria in the PCL-R manual on a 0 (does not apply), 1 (applies somewhat) or 2 (definitely applies) point scale. The resulting summary score ranges from 0 to 40. The cutoff for identifying the psychopath is typically 30. Using this cutoff approximately 15 - 20% of the prison population or .5 - 1% of the general population meet criteria for the disorder. Thus, the current assessment procedure used to identify psychopaths is comprehensive, has good reliability and validity, and importantly, has good predictive utility vis-à-vis future violent behavior (Rice & Harris, in press). Indeed, there are over 2300 citations in the scientific literature using the PCL-R.

Analyses of the PCL-R items have revealed two correlated dimensions or factors (Harpur, Hakstian, & Hare, 1988; 1989). Factor 1 includes items related to emotional and interpersonal relationships, such as superficial charm, egocentricity, grandiosity, deceitfulness and manipulativeness, and absence of remorse, guilt, or empathy. Factor 2 items reflect impulsive and antisocial behaviors, like impulsivity, poor behavioral controls, proneness to boredom, poor life planning, and irresponsibility. The second factor is the one most closely related to the Diagnostic and Statistical Manual of Mental Illness (DSM IV) classification of Antisocial Personality Disorder (APD; American Psychiatric Association, 1994).

It is important to note that although APD was intended to capture the essential components of psychopathy, it has been criticized for relying excessively on antisocial behaviors, while excluding the affective and interpersonal characteristics considered to be central to psychopathy (Widiger, et al., 1996). Thus, psychopathy and APD are not the same diagnosis.

Indeed, there is a very asymmetric relationship between the assessment of psychopathy and the criteria for APD, in that most inmates will meet criteria for APD, whereas only about 15% of inmates will meet criteria for psychopathy. In the discussion below of neuroscience experiments, studies have been limited to those in which all participants were assessed with the PCL-R in order to capture the full symptomology of psychopathy.

Psychometric tests, like the PCL-R, are designed to quantify the level of traits and behaviors within an individual. For example, if you have difficulty experiencing guilt or remorse associated with romantic relationships (e.g., you report no guilt associated with cheating on your spouse) you may score a '1' (item applies somewhat) on the PCL-R item 6 'Lack of Guilt or Remorse'. However, if you fail to experience any guilt or remorse with romantic relationships, family relationships (e.g., you steal from mom and dad), or with friends, work, school, and in criminal activities then you may merit the maximum score of 2 on item 6 (i.e., the item definition in the PCL-R manual definitely applies to you in all domains of your life). Thus, the PCL-R provides a quantitative estimate of your lack of remorse or guilt. If we find that you score high on the majority of items, you will be identified as a person who has psychopathy.

It's worth noting that self-report questionnaires are not typically considered adequate for assessing psychopathy in criminal or legal settings. Self-report instruments require

cooperation, insight, and average reading ability in those being assessed. It's also challenging for self-report questionnaires to assess traits related to pathological lying and conning/manipulative behaviors.

Another point to highlight is that psychometric tests attempt to quantify processes in the brain. Empathy, guilt, and remorse are, after all, mental processes that are instantiated in the neural systems of the brain. To understand the symptoms of psychopathy, scientists need to develop methods to accurately quantify these latter neural systems that engender the symptoms under study (e.g., lack of guilt or remorse). One common way for scientists to quantify brain processes is through the use of neuropsychological tests. These tests are largely focused on identifying problems with mental processes, such as difficulties with attention, working memory, or language. These latter processes are not typically impaired in psychopathy, so classic neuropsychological tests are of little utility in identifying individuals with psychopathy (Hart, Forth, & Hare, 1990). Future neuropsychological tests may be designed to map onto processes impaired in psychopathy, such as empathy or guilt, but at present there are no such tests in general clinical practice.

Since psychometric tests of psychopathy are proxies for measuring brain dysfunction, more direct measurement of brain processes using *in vivo* neuroimaging techniques has the potential to identify the psychopath with more accuracy than techniques that measure these processes indirectly. Let us now turn to studies that have applied direct measures of brain function to the identification of individuals with psychopathy.

Up to now, there have been approximately sixteen neuroimaging studies of psychopathy. Eight study brain structure, and eight study brain function, using functional MRI or single photon emission computed tomography (SPECT). The main goal of these studies is to relate the PCL-R score to measures of brain structure or function. These studies have largely supported the hypothesis that the emotional regions of the brain, also referred to as the limbic or paralimbic system, are disordered in psychopathy, and they appear to be disordered from a very early age. The regions most commonly implicated in psychopathy include the amygdala, orbital frontal cortex, anterior cingulate, and anterior temporal cortex.

Unfortunately, there are a number of limitations within these neuroimaging studies that hamper firm conclusions, including small sample sizes, analyses limited to only some regions of the brain, and, importantly, varying cutoff criteria for psychopathy and limited

numbers of individuals scoring over 30 on the PCL-R. Indeed, the sixteen published neuroimaging studies in psychopathy seem to include only about 105 individuals with PCL-R scores over 30 (see review by Harenski, Hare, & Kiehl, in press). These limitations make it difficult to draw definitive conclusions regarding the ability of neuroimaging to identify individuals with psychopathy.

Along with neuroimaging measures such as MRI or SPECT/PET, researchers have also used measures of brain electrical activity, such as electroencephalography and event-related potentials to study psychopathy. These latter neuroscience techniques provide limited information about where in the brain the effects are generated, but provide excellent information about the timing—on the order of milliseconds—of when abnormalities are observed.

Numerous studies have shown that psychopathy is associated with aberrant brain electrical responses following processing of demanding attentional stimuli, emotional word stimuli, and concrete or abstract word stimuli (Kiehl, 2006). Indeed, the presence of these abnormal brain electrical responses was nearly diagnostic of psychopathy (Kiehl, Bates, Laurens, Hare, & Liddle, 2006). Forty of forty-one criminal psychopaths (defined as a PCL-R score of 30 or above) were characterized by the presence of the aberrant brain electrical response, but none of the forty nonpsychopathic individuals was found to have the response (Kiehl, et al., 2006).

However, it is now known that this abnormal electrical response is also found in individuals who have brain damage to the medial temporal lobe, including the anterior temporal cortex and amygdala (see Kiehl, 2006 for a review). Thus, the abnormal electrical response in Kiehl et al. (2006) was not diagnostic of psychopathy, but rather appeared to be diagnostic of abnormalities in medial temporal lobe structures. It is worth noting that patients with brain damage to medial temporal lobe experience some symptoms of psychopathy (Kiehl, 2006).

In conjunction with other converging lines of scientific evidence, a strong argument can be made for the presence of abnormalities in limbic brain systems in psychopathy. However, research still needs to clarify the specificity of these deficits, their origin and stability over the lifespan, and their diagnostic utility. Thus, we are not currently at the point where we can use neuroscience to definitively identify, or diagnose, individuals with psychopathy.

The neuroscience of psychopathy is rapidly developing. We expect that within a few years the field of psychopathy will parallel other, more mature, research fields that have made

excellent progress using neuroimaging techniques as diagnostic tools. For example, since the time of Kraeplin (Kraepelin, 1919) and Bleuler (Bleuler, 1911) psychiatry has struggled to differentially diagnose patients with psychosis (e.g., hallucinations and delusions) into schizophrenia or bipolar diagnosis. Indeed, it currently takes six months from first psychotic episode to differentially diagnose schizophrenia and bipolar illnesses according to the DSM IV.

Neuroimaging measures can now be used to distinguish schizophrenia from healthy controls with ~95% accuracy (Calhoun, Kiehl, Liddle, & Pearlson, 2004). More importantly, functional neuroimaging measures have been shown to differentiate schizophrenia from psychotic bipolar disorder with ~90% specificity (Calhoun, Maciejewski, Pearlson, & Kiehl, 2008). Thus, the tools and techniques required to help diagnosis major mental illnesses like schizophrenia, bipolar disorder, and psychopathy are within reach of neuroscience—it is only a matter of time.

Table 1. 20 items of the Hare Psychopathy Checklist-Revised (Hare, 1991; 2003). Items from the two factor model of psychopathy are listed (Harpur et al., 1988, 1989). The two factor model labels are Interpersonal/ Affective (Factor 1) and Lifestyle/Antisocial (Factor 2). Items with '-' did not load on any factor.

	Item	2 Factor Model
1	Glibness/Superficial Charm	1
2	Grandiose Sense of Self Worth	1
3	Need for Stimulation	2
4	Pathological Lying	1
5	Conning/Manipulative	1
6	Lack of Remorse or Guilt	1
7	Shallow Affect	1
8	Callous/Lack of Empathy	1
9	Parasitic Lifestyle	2
10	Poor Behavioral Controls	2
11	Promiscuous Sexual Behavior	--
12	Early Behavioral Problems	2
13	Lack of Realistic Goals	2
14	Impulsivity	2
15	Irresponsibility	2
16	Failure to Accept Responsibility	1
17	Many Marital Relationships	--

18	Juvenile Delinquency	2
19	Revocation of Release	2
20	Criminal Versatility	--

References

Bleuler, E. (1911). *Dementia Praecox or the Group of Schizophrenias* (J. Zinkin, Trans.). New York: International Universities Press.

Calhoun, V. D., Kiehl, K. A., Liddle, P. F., & Pearlson, G. D. (2004). Aberrant localization of synchronous hemodynamic activity in auditory cortex reliably characterizes schizophrenia. *Biological Psychiatry*, *55*, 842-849.

Calhoun, V. D., Maciejewski, P. K., Pearlson, G. D., & Kiehl, K. A. (2008). Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. *Human Brain Mapping*, *29*, 1265-1275.

Cleckley, H. (1941). *The mask of sanity* (1st ed.). St. Louis, MO: Mosby.

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences*, *1*, 111-119.

Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist-Revised*. Toronto: Multi-Health Systems.

Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist-Revised* (2nd ed.). Toronto: Multi-Health Systems.

Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the Psychopathy Checklist. *Journal of Consulting and Clinical Psychology*, *56*, 741-747.

Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment*, *1*, 6-17.

Hart, S. D., Forth, A. E., & Hare, R. D. (1990). Performance of criminal psychopaths on selected neuropsychological tests. *Journal of Abnormal Psychology*, *99*, 374-379.

Kiehl, K. A. (2006). A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. *Psychiatry Research*, *142*, 107-128.

Kiehl, K. A., Bates, A. T., Laurens, K. R., Hare, R. D., & Liddle, P. F. (2006). Brain potentials implicate temporal lobe abnormalities in criminal psychopaths. *Journal of Abnormal Psychology*, *115*, 443-453.

Kraepelin, E. (1919). *Dementia Praecox and Paraphrenia* (R. M. Barclay, Trans.). New York: Kreiger.

Rice & Harris (in press). Psychopathy and violent recidivism. In Kiehl, K.A and Sinnott-Armstrong (Eds), *Psychopathy and Law*. Oxford University Press.

Widiger, T. A., Cadoret, R., Hare, R., Robins, L., Rutherford, M., Zanarini, M.,...Frances, A. (1996). DSM--IV antisocial personality disorder field trial. *Journal of Abnormal Psychology*, *105*, 3-16.

Has neuroscience already appeared in the courtroom?

Scott T. Grafton, M.D.

Most judges will already be familiar with neuroscience evidence in the courtroom that is based on expert opinion from neurologists or psychiatrists. This clinical neuroscience is usually presented in a medical framework, with a clinician integrating the history, physical examination and routine laboratory tests to diagnose a plaintiff or defendant.

Additionally, other sources of neuroscience might be seen, particularly clinical imaging and electrophysiology which can be used in this process to corroborate clinical findings and screen for pathology.

While considering the evolution of neuroscience evidence in the courtroom, it is useful to distinguish traditional clinical methods based on routine diagnostic imaging from newer imaging methods that are rarely used for clinical diagnosis.

When imaging is used to demonstrate damage to a victim's nervous system, a judge can expect that most evidence will be based on conventional clinical procedures. In cases involving victims, conventional CT scanning has dominated as the preferred diagnostic method. Less frequently, conventional X-rays, diagnostic MRI scans, electroencephalography and SPECT scanning of brain blood perfusion are used (Wolf & Stensvad, 2010).

All of these measures are usually acquired in hospitals or clinics under conditions where many technical requirements that guarantee quality assurance must be met (otherwise the sites would not be able to gain reimbursement from insurance or Medicare). All these methods are widely available. For example, over 72 million CT scans were performed in the United States in 2007. They are based on decades of procedural refinement with many expert neurologists and radiologists interpreting this information.

With this level of experience nationwide it would be unusual to hear a legal argument attempting to prevent their introduction into the courtroom based on technical grounds. Instead, courtroom arguments related to these conventional methods usually center on differences of interpretations from experts as they read a scan and consider the diagnostic specificity (does a scan, for example, show a tumor, scar or infection), or debate the relevance of the scan findings in relationship to the other clinical information at hand.

In cases where neuroscience is introduced on behalf of the defendant, a judge can expect a far more heterogeneous set of arguments and more varied methodologies, not all of which reach the level of standardization typically found with the methods involving victims (Wolf & Stensvad, 2010).

Circumstances where there are attempts to introduce some form of imaging or physiological assessment include mitigation during sentencing, demonstrating ineffective assistance of counsel presented during guilt or sentencing phase because a scan was *not* done, trial error or abuse of discretion for *not* obtaining scans, establishing insanity, diminished capacity or incompetence at some point in trial.

In these circumstances, a judge is more likely to be faced with evidence that is not part of the standard clinical repertoire used by neurologists, psychiatrists or radiologists. Because of the relatively weaker evidentiary rules for introducing mitigating evidence in capital cases, this area often forms a front line for introducing new or investigational measures of brain structure or function into the courtroom. Three important examples stand out: quantitative EEG, PET scans and fMRI scans.

Quantitative EEG emerged in the 1980s as a computer based method for analyzing electroencephalographs (EEG). In traditional EEG, a 10+ minute recording of brain electrical activity, measured on the scalp, is visually inspected by an expert reader. The reading is usually performed by neurologists who have completed additional fellowship training with this method. The examiner looks for focal or global abnormalities in electrical activity as well as sporadic abnormalities indicative of a possible source for epilepsy.

In quantitative EEG (qEEG), a digitized recording of the EEG is passed through a series of signal processing algorithms. These are mathematical transformations of the data that attempt to classify the data into a normal or abnormal category based on frequencies, events or localization. The quantitative computer analysis is particularly useful when there are many hours of data to sift through. The method is most commonly used (and paid for by insurance carriers) in clinical settings requiring prolonged monitoring in intensive care, for intraoperative monitoring, and when working up a patient up for epilepsy when protracted monitoring is required to detect where a seizure might originate. It has also been useful for identifying slow brain waves associated with dementia.

In the 1990s attempts were also made to identify electrical signatures in qEEG that could diagnose psychiatric and behavioral disorders among others. This led to a strong criticism of

the qEEG method when it became clear that the specificity and sensitivity for diagnosis in these disorders was inadequate to justify their general use (Nuwer, 1997).

The American Academy of Neurology and American Clinical Neurophysiology Society determined that the validity of these methods was *insufficient* for diagnosis of any of the following conditions: post-concussion syndrome, mild or moderate head injury, learning disability, attention disorders, schizophrenia, depression, alcoholism, drug abuse, tinnitus, predicting response to psychotropic medication, and insomnia. Diagnosis of these and other behavioral disorders by qEEG is considered to be investigational research. Furthermore, there are no standardized procedures or methods for interpreting this data and few are trained in the methodology.

Positron emission tomography (PET) scanning involves the injection of radiotracers into the blood. These circulate to the brain where the distribution of radioactivity is measured with a sophisticated ring of detectors. Depending on what is injected, the radioactivity distribution can be related to blood flow, glucose consumption, protein synthesis, drug binding, normal neurotransmitters or their synthesis, cancer and other biomolecular processes (Czernin & Phelps, 2002).

In the 1980s PET imaging of brain metabolism appeared in the courtroom to determine whether patients might have abnormalities that could not otherwise be visualized by CT scan or diagnostic MRI. It became clear that some clinical diseases, such as Alzheimer's disease and other forms of dementia had metabolic abnormalities in distinct patterns that were sufficiently specific that PET could be used as a diagnostic tool (Van Heertum *et al*., 2004). Less clear has been the clinical utility of PET metabolic imaging for diagnosing psychiatric or behavioral disorders (Mayberg, 1996). Results in mild traumatic brain injury have also been inconsistent (Granacher, 2008).

Many technical steps must be met for a PET scan to be acquired reliably, including radiotracer synthesis, blood sampling procedures, consideration of labeled metabolites, and processing of the digital scans. The method is not widely available, limiting expertise in the use of this method to a modest number of practitioners. Given these complexities, an editorial in The Journal of Nuclear Medicine in the early 1990s cautioned against the use of PET metabolic imaging as mitigating evidence (Mayberg, 1992).

Functional magnetic resonance imaging (fMRI) is a method developed in the mid 1990s to indirectly measure brain blood flow. Blood flow is coupled to brain activity. By taking scans

during a task, such as moving the hand, versus scans during no task, it is possible to perform a digital subtraction of the scans and find brain areas more active in association with the particular task.

In practice, many scans are acquired and a statistical map of the task effects is computed. These scans are just beginning to be used clinically to demonstrate a lack of brain activation in persons with brain trauma, dementia, stroke and depression. They are also used by neurosurgeons to find and avoid motor, speech or sensory areas as part of their pre-surgical planning.

In extending this to legal settings, the general strategy has been to show that the brain is not "turning on" during a particular task in a normal manner. However, this is a difficult assertion to confirm, both technically and conceptually. Diagnosticians use many different kinds of tasks on many different kinds fMRI scanners and with many different analysis methods, making it difficult to interpret findings obtained from different imaging centers.

The method is noisy, inconsistent from scanner to scanner, variable from day to day, and extremely sensitive to artifacts when someone moves their head. Thus, the scanner must be of very high quality and the subject must be exceptionally cooperative. Furthermore, establishing that a brain area is over- or under-active for a given task requires that the subject be compared to a relevant comparison group of "normal" subjects doing the same task in the same way.

Conceptually, this begs the question of what is "normal," particularly when one discovers that brain activation depends on gender, age, handedness, cooperation, alertness, level of caffeine, menstrual cycle, cerebrovascular disease, medications, recreational drugs, and strategies used to perform a task (Jezzard & Song, 1996).

A common feature of qEEG, PET and fMRI is that they are all based on sophisticated signal processing algorithms that generate some form of statistical map that is describing what the brain might be “doing” during the testing. These maps can be generated with many software methods and they invariably require some form of statistical threshold.

This leads to a fundamental source of operator bias. A small decision about how to analyze the data or where to set a threshold of what is statistically significant or not can have a large impact on whether an abnormality is present or not. Unlike a regular X-ray, CT or MRI, the

clinician does not simply look at the film and render an opinion. Instead, a statistical map is compared to some sort of reference population and an indirect opinion is offered.

The impact of these issues for the judge is that there will be a greater likelihood for arguments about the scanner, acquisition and methods used to make the results: Whether the data was obtained correctly, analyzed with the proper software, compared to an appropriate reference population and so on. It may be possible to reject this evidence on technical grounds alone. Technical experts are not necessarily the same as experts in clinical diagnosis, and testimony concerning clinical diagnosis based on the method is irrelevant until the technical limitations are first surmounted.

A judge might be faced with the argument that since there are already papers published in peer review journals using similar scanners or tasks then the method is scientifically accepted and therefore admissible. However, the results of a scientific study for one specific purpose do not readily translate to legal-diagnostic settings. The methods in a published study usually don't equate with the measurement conditions that were used to obtain data for a particular court case. There is enormous variation in scanners, software, tasks and expertise.

Thus, a judge cannot avoid addressing issues of technical competence, reproducibility, standardization and relevance when considering whether these methods should be admitted.

Assuming that neuroscience evidence is admissible from a technical standpoint, there still remains the issue of whether the information will actually matter to the case in hand (Garland & Glimcher, 2006). While any of these methods hold the potential for demonstrating an abnormality of the brain, the legal relevance may not be clear (Batts, 2009). A finding may be nonspecific or incidental. Whether it was present at the time of a crime may be unknown. The causal link between a scan abnormality and an individual's behavior is a complex and evolving question for research.

Finally, courts must consider whether, in contrast to the attempted use of scans to mitigate punishment or reduce responsibility, scans might alternatively be used to attempt to show increased responsibility or indicate additional punishment is required.

References

Batts, S. (2009). Brain lesions and their implications in criminal responsibility. *Behavioral Science and the Law*, *27*, 261-272.

Czernin, J., & Phelps, M. E. (2002). Positron emission tomography scanning: current and future applications. *Annual Review of Medicine*, *53*, 89-112.

Garland, B., & Glimcher, P. W. (2006). Cognitive neuroscience and the law. *Current Opinion in Neurobiology*, *16*, 130-134.

Granacher, R. P., Jr. (2008). Commentary: Applications of functional neuroimaging to civil litigation of mild traumatic brain injury. *Journal of American Academy of Psychiatry and Law*, *36*, 323-328.

Jezzard, P., & Song, A. W. (1996). Technical foundations and pitfalls of clinical fMRI. *Neuroimage,4*, S63-S75.

Mayberg, H. S. (1992). Functional brain scans as evidence in criminal court: an argument for caution. *Journal of Nuclear Medicine*, *33*, 18N-19N, 25N.

Mayberg, H. S. (1996) Medical-legal inferences from functional neuroimaging Evidence. *Seminars in Clinical Neuropsychiatry*, *1*, 195-201.

Nuwer, M. (1997). Assessment of digital EEG, quantitative EEG, and EEG brain mapping: Report of the American Academy of Neurology and the American Clinical Neurophysiology Society. *Neurology*, *49*, 277-292.

Van Heertum, R. L., Greenstein, E. A., & Tikofsky, R. S. (2004). 2-deoxy-fluorglucose-positron emission tomography imaging of the brain: current clinical applications with emphasis on the dementias. *Seminars in Nuclear Medicine*, *34*, 300-312.

Wolf, S. M., & Stensvad, E. (2010). Neuroscience in reported criminal law decisions: 1994-2009. In: *MacArthur Foundation Program in Law & Neuroscience*. Los Angeles, CA.

How is neuroscience likely to impact law in the near future?

Read Montague, Ph.D.

Towards the end of the twentieth century, neuroscience emerged as a vast, multifaceted discipline. Although there are many reasons behind this growth, one driving force has been the widespread development of new physical techniques for probing biological function at many time and space scales.

These techniques include new physical probes for cellular and molecular function mainly applicable in animal models, but they also include computational methods applied to everything from molecular modeling to large-scale modeling of cognitive function in humans. In addition to this remarkable progress, one of the most exciting developments over the last fifteen years has been the capacity to image the structure and function of living human brains.

There is now a collection of imaging and computational methods that allow scientists to probe healthy human brain function and many of them have been described in the preceding chapters. These include CT (computed tomography), MRI (magnetic resonance imaging), fMRI (functional magnetic resonance imaging), EEG (electroencephalography), MEG (magnetoencephalography), and the nuclear imaging methods PET (positron emission tomography) and SPECT (single photon emission computed tomography). Various combinations of these imaging modalities are also now available *(e.g.*, PET-CT). Each method is good at imaging different kinds of structure and function, but the upshot is that we now have an active and growing treasure trove of means to look directly past the skull at the structure and function of healthy human brains.

Connecting mind to brain: some limiting features

Our current best efforts to relate mental function to brain activity all depend on one fact that will be part of human neuroscience into the foreseeable future: the connection between cognitive variables and any physical measurements in the brain will not be deterministic, but will result from statistical inference. This is not surprising; our best experimental evidence in physics is subject to the same kind of data-dependent limitations.

Now on occasion that connection between brain signals and mental function may turn out to be very regular and will possess a high degree of certainty, but the connection will always have an element of uncertainty. This is true because of the nature of real-world data, and it's truer for a biological device like the brain, because variability is the lifeblood of all living things – it is the basis for natural selection. So we should expect a lot of variability within and between brains and the data we have on this issue shows this to be true.

But there is an extra bit to the problem: the variability in mental function and the brain function that underlies it has really never been characterized experimentally. This systematic lack of data poses a real impediment to the useful contribution of neuroscience data to legal issues.

While there exist admirable efforts to quantify structural variability in health, disease, and even through development, there is no large-scale effort to quantify variability of functional imaging data with detailed variability in measurable cognitive variables. This is a crucially underdeveloped area of human neuroscience and easy to understand.

Functional imaging techniques are young and our experience with how to produce such "at scale" depictions of human mental function is just beginning. Nevertheless, these kinds of databases would need to exist to make statistical statements about the normal range of variation along some brain or behavioral measure.

In order to make any statements about the capacities (or lack therefore) in an individual, one would need to quantify the distribution of human variation along some cognitive dimension and the connection of such variation to a similar distribution along some brain measurement. Variation in healthy human cognitive function has not generally been connected to variations in healthy human brain function. Notice that in all this discussion, neither the exact nature of the physical measurement nor the cognitive variable matters for the strength of the argument.

This gap in our knowledge could be reasonably closed in the near future by larger-scale efforts to relate normal variation in brain function to normal variation in cognitive function. This would simply require a choice of which cognitive dimensions to measure and a plan for running a sufficient number of subjects so that entire distributions of responses could be quantified. Such a database would be valuable not only to the legal system, but to cognitive neuroscience more broadly.

How should we compare individuals to distributions?

The limitation outlined above relates to the quantification of variability in populations of humans. As mentioned, a concerted effort in the near-term future could remedy the situation significantly. However, even in the presence of a solution to the population problem, there is a more profound limitation that relates to the lack of data for individuals along with the lack of accepted methods for how to compare an individual's (brain and behavior) to a distribution of data drawn from a population.

The comparison problem possesses two major sub-problems – (1) comparing an individual's brain and behavioral responses now (in the present) to a distribution of normal healthy responses, and (2) using some measurement now of an individual combined with other information to retrodict what that individual's likely state was "at the time of some prior event" (like a crime).

While the law has some standards for addressing elements of these two issues, neuroscience does not. While there are acceptable statistical methods available to compare an individual response to a measured distribution of responses, in the case of human brain and behavioral responses, we do not now possess standards for doing this systematically.

We can take a concrete example concerning risk perception since it will highlight clearly the gaps in our knowledge. The example concerns the prospect of carrying out some act (say driving a car fast) that carries an externally measurable risk of injury to self and others. By "externally measurable" I mean possessing an objective procedure for deciding the level of risk along some prearranged scale. For this example, we will separate the problem into two parts: (1) the subjective perception of the objective risk by an individual, and (2) the capacity in that individual to act on their own risk perception.

Let us avoid assigning any legal valence to these issues here, but use this decomposition to expose our knowledge gaps. Below we outline the kinds of scientific probes or knowledge that would be needed.

1. The individual's subjective perception of objective risk.

To measure this quantity, we would require some kind of independent neural signature that reflected the subject's perception of the objective risk. By "independent" we mean independent of any actions taken by the subject – that is, we want to avoid using observed actions by the subject to infer the subjectively perceived risk. Current experimental

paradigms would ask a subject to select from a series of risky choices and then attempt to make an inference (based on the observed choice) about the risk perceived by the subject.

The problem with this method is that it conflates two issues (and perhaps three) – we cannot tell the difference between perceived risk, capacity to act on that risk, and, contingent on both perceived risk and capacity, the willingness to act in a certain way. We simply cannot use action-based readouts to assess an individual's subjective risk perception.

We need something like a covert neural signature that scales with an individual's perception of risk. This would require new knowledge before it could be applied.

2. **The measured capacity of an individual to act on perceived risk.**

Once some kind of independent risk perception correlate for the subject was established, we would have to establish the relationship between this perceived risk and the subject's capacity to act conditioned on that risk.

It's altogether possible that some subjects would display a kind of "risk blindness" across certain levels of perceived risk. For example, Hsu and colleagues (Hsu *et al*., 2005) have shown that subjects with orbitofrontal cortex lesions lack a sensitivity to risk – exactly the kind of risk blindness mentioned above. Now these patients are missing cerebral cortical tissue and it's obvious; however, imagine that a more subtle physiological difference existed in this same region in other subjects not explicitly deemed impaired but who exhibited a kind of blindness to risk.

The implication is that the physical instantiation of such risk blindness, whether obvious or subtle, should not matter when considering how this risk blindness affects the law.

So again, the capacity to act on perceived risk would require a new kind of assay – one where we have an assurance that the perceived risk is one value (through an independent neural signature for example), and where we can then measure the range of capacity to act given this risk.

The capacity to perceive risk and act based on it could be separate systems in the brain and might be independently subject to damage through disease or injury. Given a known level of subjectively perceived risk, we would need a way to assess the flexibility of an individual's capacity to act.

In the case of some action where injury or damage to property may occur, the task would need to assess the range of riskiness of action available to the individual in the face of a level of perceived risk. Characterizations of such ranges within an individual would be very useful, but only if the distribution of these ranges for a population of humans was also available.

Again, we need populations of brain and behavioral responses, specific and new assays in individuals, and new methods for comparing the two.

Where is the near-term payoff likely?

We must emphasize that the bulk of the argument above focused on missing data that neuroimaging paired with behavioral probes could in principle answer in the coming near-term years.

The lack of data focused on two missing areas: (1) lack of measures across populations of humans, and (2) the lack of methods and efforts to compare individual responses (brain or behavior) to the distributions yet-to-be-characterized. These missing data fall short of a prescription for how to use them to reason about the law.

Such legally balanced judgments must await the collection of the information first in order to see whether there are any unexpected surprises.

Lastly, we reiterate one important exit issue, that is, the exact nature of the brain measurements and their mechanistic origins isn't pertinent for legal issues at this point in the evolution of modern neuroscience. The important need is for consistent informative measures across populations of humans and methods for comparing individual responses to populations.

In the near-term, we think that mitigating neural deficits will most likely be restricted to individuals with demonstrable brain damage known to interfere with characterized mental function. Anything more subtle than this awaits a new style of data collection (very large scale) and a coherent effort to understand the empirical and theoretical developments required to relate an individual's brain and behavioral responses to a population, but in a manner relevant to legal issues. These efforts will require a close collaboration between legal scholars, neuroscientists, and computer scientists.

Reference

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*, 1680 – 1683.

How is neuroscience likely to impact the law in the long run?

Adina Roskies, Ph.D.

Many of the most significant technological advances in brain science thus far were unforeseeable a decade or two before, as were the advances in factual and theoretical understanding they made possible. Additionally, many predictions about the extent of future knowledge have failed to materialize. This track record suggests that the long-term influence of neuroscience on the law will be exceedingly difficult to predict with any confidence. My response therefore will be highly speculative, and the more dramatic the consequences for the law, the less credence one should have in the answer.

With that caveat, I envision three main ways in which neuroscience is poised to have a significant affect.

First, advances in neuroscientific understanding of the brain bases of behavior coupled with improved neurotechnologies will improve assessment techniques that will have an impact on diagnoses of brain disorders correlated with social misconduct and on prediction of offenses and recidivism. Second, this knowledge will enable the development of treatments for certain conditions. This will have an impact on the law both in terms of affecting the demographics of offenders, and prompting us to address certain legal questions with new policy implications.

Third, and most speculatively, neuroscience may alter commonsense views about human nature, and in so doing affect policy, and indirectly, the law.

In what follows I discuss these points concerning particular topics relevant to the law.

Lie detection and mind reading

Polygraphy is not currently admissible as evidence in most jurisdictions. The rationale for this inadmissibility is predicated largely on its inaccuracy (Council, 2003). The prospect of more reliable methods of lie detection stemming from brain research is plausible, since tapping into brain signals would in theory enable one to probe the actual mechanisms of deceptive behavior, rather than their downstream consequences. Indeed, neuroimaging methods, including event-related potentials (ERPs) and functional magnetic resonance

imaging (fMRI), are currently being investigated for this purpose, and a few companies are already marketing products to government and private citizens (Greely & Illes 2007).

It remains an open question whether there are neural signatures that reliably correlate with individual acts of deception, and that are distinguishable from other neural signals. Based on current data, it is reasonable to expect that neurotechnologies will surpass the average human, and even trained humans, in detection of deception.

A recent study suggests that fMRI is unreliable for detecting objective truth, though it is considerably better at detecting intentional deception (Rissman *et al.*, submitted). If so, these methods may be sufficiently accurate for routine use for some legal purposes, for example for exculpation or reduced sentencing.

Reliability standards for incrimination should be much more demanding than for exculpation, and it is possible that lie detection neurotechnologies will never be able to meet such standards. But they might. Achieving this level of accuracy will require a radical change in the quality of empirical studies assessing accuracy, as well as increased ecological validity.

To date, almost all experiments professing to assess the validity of lie-detection techniques have used contrived situations in which subjects are told to lie. Since following an experimenter's directions to lie is a radically different action than real criminal deception, these experiments are insufficient basis for adoption in legal settings. Additionally, these techniques will have to be applicable to singular events. Current methods require averaging results over many trials, which renders them inapplicable to most legal settings.

Neuroscience may also provide other ways for assessing truth (such as distinguishing experiential from false memory), or for determining with some degree of specificity the contents of thoughts. Although recent studies claim to enable researchers to mind-read (Mitchell *et al.*, 2008), these experiments involve discriminating among a constrained set of options.

Apart from questions of admissibility, improvements in such techniques may require us to rethink whether neuroimaging results should be considered a form of testimony or a form of physical evidence, and whether requiring them in legal proceedings falls afoul of the fifth amendment, and when their use can be compelled or covert (Stoller & Wolpe, 2007). Privacy law is apt to become important with respect to these technologies.

Addiction

Drug addiction is one of the largest problems facing the criminal justice system. Neuroscience of drug addiction is an active field of research, and it promises to have huge policy implications.

Neuroscientists are discovering the biological bases of drug action on the nervous system, and are beginning to understand the way drugs act on mechanisms of reward and learning to cement addictive behaviors. Some advances in neuroscience, such as distinguishing pathways for liking and wanting (Berridge, 2003), shed new light on addiction and may alter the way it is viewed in society.

Public views of addiction may change when people come to appreciate the difference between wanting something because it produces pleasure, and craving it because it hijacks the brain's reward systems, even when it has long ceased to provide pleasure. A change in perception of addiction may buttress efforts to revamp drug laws and to devote more energies to treatment than to punishment.

The increased understanding of the neural pathways underlying various addictions will also likely enable much improved treatments for addiction, including specific targeting of various drug pathways in the brain, and development of novel substances for blocking drug receptors or reward mechanisms (Koob *et al.*, 2009).

It seems likely that drug offenders will in the future be treated by the legal system separately from other types of offenders. Neuroscience may also significantly affect the way drug courts operate, influencing the existence of legal mandates for treatment, *etc*. While drug decriminalization is extremely unlikely, it is a conceivable policy outcome of a fully developed neuroscience of addiction, and it would have far-reaching consequences for society and the law.

Mental illness and psychopathy

The law currently has some provisions that affect the mentally ill, including avenues for mitigating punishment in capital offenses, competency requirements, and civil commitment. However, determination of mental illness is rarely made based on neuroscientific data, since neuroscience currently lacks diagnostic tools for mental illness. This is bound to change, as we increasingly discover the biological bases of a variety of mental illnesses.

It is very likely that chronic mental illnesses will be diagnosable with neuroimaging and other neuroscientific techniques, and it may even be possible to make behavioral risk assessments based on these data. It is possible that episodes of fulminant mental illness (those that arise suddenly) may be foreseeable. Such data will undoubtedly be introduced at trial for the purposes of conviction or exculpation. It is likely that the law will alter to accommodate ramifications of the availability of this kind of knowledge.

Psychopathy is a continuing problem for the legal system, because of the high rates of recidivism demonstrated by psychopaths. Although the criteria used for determining psychopathy are clinical and not neurological, evidence is mounting for a common neurological basis for psychopathic traits (Kiehl, 2006). If such data bear out, neuroscience may enable us to distinguish psychopathic from nonpsychopathic inmates, and will likely affect both sentencing and parole.

It is possible that greater understanding and predictive power will encourage the development of alternative ways to protect society from these individuals. What will have more of an impact, however, is that we will in theory be able to determine those with psychopathic traits in the nonincarcerted population. The degree to which such information can be legally acquired will have to be regulated, as will any consequences that stem from that knowledge. Until now, we have been unable to predict with any degree of success behaviors which have serious detrimental consequences for society.

Although our legal system does not allow punishment for offenses not yet committed, the law will have to take seriously the question of what actions can be taken to protect society from crimes that have not yet occurred, should such prediction become possible. One area in which we may see this arise is with counter-terrorism efforts, both domestic and foreign. Moreover, it is likely that international law will have to address differences among attitudes regarding the use of predictive power in different legal systems.

Often, along with understanding comes the means of intervention. It is likely that for many of the mental abnormalities which neuroscience will enable us to diagnose, it will also provide novel avenues for treatment. While this will likely have a positive effect on the prison system, it will raise a number of issues regarding the legality of mandating treatment, treatment distribution and enforcement. It may also raise deeper issues about the state's interference in a citizen's personhood and identity. That may not be as much of an issue

regarding drug addiction as it promises to be when the treatment affects things as deeply ingrained as one's personality, temperament, beliefs and values.

Responsibility

The preceding section raises issues that can be generalized to the population at large. It is inevitable that there will need to be a broad social debate about the meaning of responsibility in a world where the physical basis of behavior is in large part understood. Folk conceptions of responsibility are closely tied to those of freedom, and folk notions of freedom tend to be anti-physicalist (admitting of some kind of nonphysical causal influences), or at least contracausal in nature.

A number of scientists and philosophers have worried that science will reveal free will to be an illusion, and that morality and responsibility will fall with freedom. Although I think such fears are overstated, I do think that a change in view about human nature is bound to accompany a significantly developed neuroscience, and we ought to recognize that the law relies heavily upon folk notions. If the folk notions change dramatically, the law will likely follow suit.

Thus, it is by putting pressure on our intuitive notions of responsibility and agency that neuroscience may have the most far-reaching consequences for the law. This may be so although the law as currently conceived does not consider causal mechanism to be relevant (Morse, 1994). I reiterate that such an outcome is much less likely than the more concrete changes discussed above.

Currently the law operates with notions of personhood and agency that take seriously concepts of volition, control, choice, belief, desire, and responsibility. It is possible that neuroscientific advances will require revisions in some of these views, or less likely, a rejection of some of them as applicable to humans. If such notions become widely accepted, pressure will be put on the legal system to adapt to this new framework.

It is highly unlikely that moral or even legal responsibility will be threatened in its entirety. Indeed, recent experimental results suggest that people are more apt to give up other central beliefs than they are to jettison practices of blame and holding people responsible for their actions (Roskies & Nichols, 2008). However, it is likely that certain types of actions that we now treat as voluntary will be recognized as compelled, and that parts of the law will become more instrumental and less retributive in nature (Greene & Cohen, 2004).

References

Berridge, R. A. (2003). Addiction. *Annual Review of Psychology*, *54*, 25-53.

Council CtRtSEotPNR. (2003). The Polygraph and Lie Detection: National Academies Press.

Greely, H. T., & Illes J. 2007. Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law and Medicine, 33*, 377-431.

Greene J., & Cohen J. D. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London B*, *359*, 1775-1785.

Kiehl, K. (2006). A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. *Psychiatry Research*, *142*, 107-128.

Koob, G. F., Kenneth, L. G., & Mason, B. J. (2009). Development of pharmacotherapies for drug addiction: A Rosetta Stone approach. *Nature Reviews: Drug Discovery*, *8*, 500-515.

Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science, 320*, 1191-1195.

Morse, S. J. (1994). Culpability and control. *University of Pennsylvania Law Review, 142*, 1587-1660.

Roskies, A. L., & Nichols, S. (2008). Bringing moral responsibility down to earth. J*ournal of Philosophy*, 105, 371-388.

Stoller, S. E., & Wolpe, P. R. (2007). Emerging neurotechnologies for lie detection and the fifth amendment. *American Journal of Law and Medicine*, *33*, 359-375.