Neuroimaging and Criminal Law

84

Reinhard Merkel

Contents

Abstract

Methods of neuroimaging have sporadically, though in recent years increasingly, occurred in legal proceedings. By now, however, it seems that they are about to enter courtrooms on a systematic basis. This poses a host of normative problems, to do, for instance, with future applications of neuroimaging to determine culpability, to test the veracity of testimony, or to predict the future dangerousness of perpetrators. The latter two, brain-based lie detection and "neuroprediction" of dangerousness, are examined in this chapter. Functional

R. Merkel
Faculty of Law, University of Hamburg, Hamburg, Germany
e-mail: reinhard.merkel@jura.uni-hamburg.de

J. Clausen, N. Levy (eds.), *Handbook of Neuroethics*,
DOI 10.1007/978-94-007-4707-4_165,

magnetic resonance imaging (fMRI) is taken as a paradigm model, and its potential impacts on criminal trials are explored. The analysis is premised on a range of basic distinctions: between (1) different phases of a criminal trial; (2) the divergent roles played by the parties to a trial, most notably prosecution and counsel, and the different evidentiary goals and burdens associated with these roles; and (3) between compulsory and consensual fMRI. It turns out that there are no good reasons to ban fMRI for lie detection or for neuroprediction from criminal proceedings entirely. Instead, it should be admitted differentially in criminal trials, viz., only for purposes of exoneration, but not of conviction, of the defendant. Substantiating arguments are expounded. In cases of preventive detention, it may even be obligatory for the state to offer chances of possibly exonerating brain imaging to perpetrators who were otherwise considered candidates for indefinite custody.

Introduction

Unless all indications are deceptive, we are at the brink of a more or less systematic introduction of various methods of neuroimaging into legal proceedings.[1] This pertains to all major areas of the law, public as well as civil and criminal law.[2] Here I will focus on criminal law and its normative foundations only. Indeed, it seems fair to say that, for a variety of reasons, none of the other legal spheres will be affected as profoundly and in a comparably vexing manner by the suggested development as the substantive and the procedural law of crimes. This needs no elaboration here. Suffice it to say that in criminal law the state deploys its sharpest sword, as it were, against its own citizens, and hence has a constitutional duty to meet the most stringent obligations to justify its measures. Broadly speaking, this is the main reason why any uncertainty concerning validity and reliability of methods of neuroimaging as well as its potential to intrude into protected spheres of its subjects is bound to cause greater irritations in the field of criminal law than in any other legal area.

Over the last decades, a whole range of brain imaging methods have been developed or, if already on hand, significantly enhanced. For the purposes of our analysis, their classification into structural and functional techniques is of particular interest. Those of the former type provide ways of demonstrating the anatomical status of the brain, whereas the latter make its activities (functioning) accessible to outside observers. Structural procedures are, for instance, CT (computed tomography) and MRI (magnet resonance imaging); functional methods are electroencephalography (EEG), positron emission tomography (PET), single-photon emission

[1]This, of course, is a very general and hence not very informative statement. Before it can be elaborated on, it needs to be differentiated into a host of distinctions pertaining to conceptual as well as factual and normative questions – matters which I will deal with in due course.

[2]Least so, presumably, to public law, for reasons not to be developed here. On the various potential applications in civil law (see Granacher 2012; Eggen and Laury 2012; Moriarty et al. 2013).

computed tomography (SPECT), and functional magnetic resonance imaging (fMRI). In what follows, I will concentrate on methods of fMRI. They are currently the most technically advanced and, in various respects, the most promising ways to explore brain functions. Hence, the normative problems of neuroimaging that we will focus on can best be expounded by examining this paradigm model.

I will not talk about the clinical merits of brain imaging, i.e., its diagnostic, therapeutic, and preventive values which are largely beyond dispute. Instead, I will investigate some of its nonclinical applications. All of them have long become objectives of extended research in basic as well as in applied sciences. But some of them have also entered the courtroom. That is what makes them normatively interesting in various respects, some of which appear to be new and unique.

Basic Science: Techniques

As indicated, magnetic resonance imaging (MRI) in its standard forms can be either structural or functional, depending on what it is destined to measure: anatomical structures of a brain or certain of its activities, viz., functions (Bigler et al. 2012). Structural images have their primary field of application in clinical contexts of diagnostic or therapeutic provenance.[3] More interesting for the purposes of criminal law as well as legal ethics is functional magnetic resonance imaging (fMRI). For according to the expectations commonly associated with its use, it is supposed to make observable to external examiners, though only in an indirect way, the neural correlates of certain mental processes in the minds of persons tested in a scanner while being engaged in various cognitive tasks or exposed to stimuli designed to evoke emotive reactions. And this is said to procure substantive ground from which the goings-on in subjects' minds become accessible (deducible) in some sense for outside observers. Sure enough, most of this wording amounts to claims that are much too sweeping. They require several conceptual and technical qualifications as well as normative *caveats*. We will turn to these later.

The physics and technical intricacies of fMRI are very complex, much more so than could be surmised by looking at what the media tell us about it, with their basic microphysical processes rooted in fundamentals of quantum mechanics. For our purposes here, we need not delve into such difficulties. The following sketchy account may suffice.[4] FMRI is an imaging technique that records certain "changes

[3]They do, of course, also have various applications in the law, e.g., in civil cases, such as medical malpractice suits where they might be used to demonstrate differences in structural brain conditions before and after an intervention, etc. Normatively, these are mostly trivial problems of no particular pertinence to imaging methods.

[4]For a brief yet still formidable description see Logothetis 2008; extensively Logothetis and Pfeuffer 2004; for overviews accessible to laypeople see Raichle 2009; Jones et al 2009; Langleben et al. 2012; Taylor 2012, pp. 103–110.

in the regional blood volume and flow that are associated with cognitive activity" (Langleben and Moriarty 2013). Oxygen consumption, glucose utilization, and therefore blood flow in the brain constantly change in ways that are closely and in a remarkably precise mode related to cellular activities in the respective neural areas (Raichle 2009, p. 118). Such cellular activities, in turn, are believed to be the substrate, and hence (in some sense) the physiological mirror, of mental processes of all kinds in the minds of subjects during their test in a scanner.

The data acquired in fMRI are based on differences in the magnetic properties of the content of blood vessels and surrounding tissues in the brain as well as on such differences between oxygenated (aortic) and deoxygenated (venous) blood. These differences are elicited and made measurable by exploiting differences in the "resonant" properties of subatomic particles (protons) of hydrogen atoms contained in the water molecules of brain cells and surrounding blood flows, namely, their varying capacities to resonate (i.e., to absorb and emit energy quanta) while being exposed to brief pulses of radiation. This is achieved by placing the head of the subject into a powerful magnetic field in the range of 0.5–5 T, generated and modified by the scanner, thus uniformly aligning all magnetic particles in the targeted brain areas, and then emitting large numbers of very brief consecutive radio waves which are also produced by the scanner. These pulses deflect the previously aligned magnetic axes of the hydrogen protons away from their state of collective equilibrium. Soon as a radio pulse stops, the protons return to their former equilibrium, thereby "resonating," i.e., emitting miniscule radio quanta that can be detected by the scanner. These signals vary in strength depending on which type of tissue the emitting protons belong to. Thus, they characterize the diverse organic matter from which they originate: oxygenated or deoxygenated blood or other tissue. It is these differential radiation echoes that the scanner records and processes in order to create its fMR images (cf. Matthews and Jezzard 2004).

The specific brain activity believed to be the substrate or (at least) correlate of certain mental activities is indicated chiefly by the varying "echoes" of either oxygenated or deoxygenated blood particles, hence the label these signals are commonly marked with: "blood oxygenation-level dependent" (BOLD). They are of particular interest to researchers or clinicians. BOLD imaging is presently the most widely used fMRI technique (Langleben and Moriarty 2013, p. 223).

The scanner also permits localizations of the collected signals in (brain) space, endowing the method with a high quality of spatial resolution.[5] By applying large quantities of radio pulses from different angles all around the head and recording

[5]Temporal resolution of fMRI is of necessity weaker since the physiological reactions of the brain to the various tasks presented to its bearer, i.e., the increased blood flow to its respective parts, take a few seconds time.

the diverse echo pulses thus elicited, huge amounts of data are collected by such successive measurement "slices" through the brain. The scanner further subdivides these slices into relevant units of volume, so-called voxels,[6] tiny cubicles of (currently) 1–3 mm per side. Data derived from the voxels can then be configurated by a computer program into 3D images of the brain. Subsequently, these images are colored differentially by the computer and projected onto a baseline picture of the individual brain, usually taken beforehand by a structural MRI. The colors used in this process accord to the varying quantities of blood flow indicated by the recorded BOLD signals. They have no inherent meaning and do not mirror a colorful reality in the brain. It is the researchers that (to some extent arbitrarily) specify what, i.e., which amount of differential neural activity, is expressed by each color. By convention, the general rule is that "the brighter the color (say yellow compared to orange) the greater the statistical significance of the differences between two conditions" (Jones et al. 2009).

It is important to note that these data do not depict any "absolute" measure of regional brain activity. Rather, what they indicate is "relative regional activity over time" during the test (Langleben and Moriarty 2013, p. 223). This is achieved by relying on a method called "cognitive subtraction" (Aguirre and D'Esposito 1999). Here is a brief expert description:

> This principle assumes that the fMRI signal difference between two behavioral conditions that are identical in all but a single variable is due to this variable. Therefore, a proper comparison (i.e., control) condition is critical for meaningful BOLD fMRI paradigm. The fMRI activation maps reported in the literature usually represent a statistical subtraction between the fMRI activity maps related to the target and control variables. Ideally, the comparison and target conditions would be identical, except for a single variable of interest (Langleben and Moriarty 2013, p. 223; see also Raichle 2009, p. 5).

Reliance on the "cognitive subtraction" method demands observance of elaborate criteria for highly sophisticated experimental designs on the part of the researchers if their endeavor is to be successful in terms of clinical relevance or meaningful in terms of basic science.

Metaphysical Problems of Mind and Brain?

This method, however, also accounts for quite another fundamental aspect of fMRI. Philosophers who have read their way through the vast literature on fMRI might wonder why there is never any mention, not even in passing, of the age-old metaphysical problem of the relationship between the mind and the brain. If one aspires to measure (albeit indirectly) the brain activity of a person in order to disclose certain mental processes in their minds, does that not imply that one adheres to a certain type of monistic materialism (or physicalism) in the said metaphysical debate, that is, to a position holding that all mental

[6]Derived from "volume element" in assonance with the two-dimensional visual unit "pixel."

events are either simply identical to ("nothing over and above") correlated processes in the brain or yet "supervene" on such processes and hence metaphysically depend on them? And is that simply to be taken for granted? In other words, aren't there any reputable dualisms around anymore in contemporary metaphysics?[7]

We may, however, leave this issue aside here as unresolved as it is in philosophy. The "cognitive subtraction" principle of fMRI makes sure that, irrespective of fundamental mind-brain metaphysics, fMRI scans are apt to deliver empirical results that can reasonably be taken as disclosures of relevant facts about the mind. For what they aim to determine is not some "absolute" mental state "in itself," plainly read off its neural substrate. That would certainly amount to a more magical type of science fiction. Rather, they only seek to establish a difference between two brain states by comparing two different patterns of brain activity. One of these is the "baseline" (or "control") state before the test person is confronted with a particular cognitive task, whereas the other is task related, i.e., induced by the particular variable in the experimental setting which is introduced by the test-specific performance. If a difference between the brain activation patterns of the baseline state and the task-related state consistently emerges while the subject performs a specific cognitive task, one can be reasonably sure that the variation in those mental states manifests itself somehow in the recorded difference in brain patterns and can thus be identified by the latter. Put another way, that the resulting "difference image" would represent only those neural areas concerned with the specific task of the particular test (Raichle 2009, p. 5). This is quite independent from whatever else the relation between mind and brain might be in terms of metaphysical monism, reductionism, dualism, supervenience, or even Cartesian causal interactionism. Hence, we can safely lay aside these unresolved problems of the philosophy of mind.

Neuroimaging of Deception: Feasibility? Admissibility?

As indicated above, our normative analysis has to start with some fundamental distinctions.

1. Firstly, one should keep apart the different purposes for which fMRI could possibly be utilized in criminal law procedures. Three of them are rather obvious

[7]For elaborated overviews on mind-brain identity theories, on supervenience and on dualism(s), see Stanford Encyclopedia of Philosophy, entries "Mind/Brain Identity Theory" (Smart 2013), "Supervenience" (McLaughlin and Bennett 2013), and "Dualism" (Robinson 2013). To be exact, supervenience theories, holding that mental processes are asymmetrically dependent on corresponding brain processes, are really a form of moderate dualism (albeit decisively different from all types of "substance dualism" in the wake of Descartes' classical position which has only few followers these days). Supervenience, of course, is perfectly compatible with the assumption that measuring certain brain processes via fMRI can disclose correlated mental events.

(albeit not exhaustive): lie detection, assessment of responsibility, and prediction of future dangerousness. In what follows, I will confine my analysis to the first and last of these issues, i.e., on neurodetection (of lies) and neuroprediction (of dangers).[8]

2. Secondly, the question whether fMRI is a sufficiently reliable method of evidence in criminal law and hence admissible in the respective legal proceedings should be distinguished from the normative question whether such applications can be justified in principle, which is to say, even if fMRI turned out to be a scientifically valid and reliable method of evidence. It is by no means clear ex ante which of the risks possibly posed by the use of neuroimaging in criminal law appears more alarming: that of their uselessness or of their potency for evidentiary purposes.
3. Thirdly, the different roles of the parties in criminal trials[9] associated with divergent, in part conflicting, interests must be taken into account. Not all of these roles are equally suited for an attempt to deploy neuroimaging methods for one's particular aims.
4. Furthermore, a criminal trial is not a homogenous process with one invariable objective (be it retribution, prevention, a mixture of both, or something else) remaining constant through all parts of the proceedings, and with immutable interests of the parties involved as well as of the state and the public. Rather, it consists of a number of distinct phases, each of which is assigned a different principal purpose by the law and in each of which the roles of the parties and of the witnessing public also change to varying degrees.
5. Finally, of particular importance here is the distinction between the trial procedures in court (criminal proceedings proper, as it were) and, as the case may be, the subsequent enforcement of a prison term. During the latter too, insights into the mental sphere of a prisoner, possibly procured by an fMRI, might well be of special importance, be it to the prisoner himself or to the state and the public, for instance, if such insights were apt to sustain a reasonable prediction about the future dangerousness of the prisoner after a potential discharge from jail.

All of the above distinctions are of a more or less rough-and-ready kind. But they must be taken into account as one reflects on the problems of fMRIs in criminal procedures. All of them are associated with varying kinds and degrees of significance that the results of a neurotechnical access to the mental sphere of a party, be it the defendant or a witness, might have for that party itself as well as for the others in

[8]The important issue of neuro-assessing guilt and responsibility would require a whole treatise of its own, presupposing a clarification of the concept of (criminal) responsibility, which in turn implies hotly contested issues of freedom of the will, including the "principle of alternative possibilities" (PAP) as an alleged prerequisite of free will and responsibility, and other perennial issues in the metaphysics of mind. I have developed my own thoughts on these topics elsewhere (cf. Merkel 2008, 2011).

[9]Mainly the accused and his or her counsel; the prosecutor, judges, and jury; witnesses and expert witnesses; and, to a limited extent, the victim of the tried criminal offence.

every phase of the process. Put in the sweeping form of whether or not neuroimaging is admissible to courts in criminal law, the question admits no sensible answer. So the following considerations will keep in line with the foregoing delineations.

Lie Detection in the Trial Phase: Validity? Reliability?

For quite some time, measurements of physiological indicators of deception, such as skin conductivity, blood pressure, respiration, and a few more, have been conducted by employing the traditional "polygraph" which measures activity in the peripheral nervous system to detect deception. Two different methods of acquiring the relevant data have been used: the "control question test" (CQT) and the "concealed information test" (CIT), or more popular "guilty knowledge test" (cf. MacLaren 2001). In the course of the former, usually three types of questions are posed to the test person: firstly, incriminatory ones (such as "Did you kill your wife?"); secondly, irrelevant and harmless control questions (such as "Who is the current president of the United States?"); and thirdly, control questions about nonspecific misconduct, designed to somehow strain even nondeceptive subjects, but not in an specifically inculpatory way, i.e., not with regard to the particular incident in question (such as "Have you ever mistreated your wife or your children?"). By contrast, the CIT takes aim at highly specific and crime-related knowledge that subjects could only possibly have if they were indeed involved in the perpetration of the criminal offence in question. If that is the case, their autonomic nervous system will respond differently from that of an innocent person to such interrogating stimuli. Or so it is claimed.

The polygraph has come into disrepute, at least with regard to its application in criminal cases. Most, though not all, courts in Europe as well as in the United States have disapproved polygraph-based evidence (cf. National Research Council 2003, Chaps. 3 and 4). This pertains particularly to its CQT version as opposed to the CIT. The latter method whose reliability appears to be significantly superior to that of CQT (Ben-Shakar and Elaad 2003, p. 132; MacLaren 2001) has also been extensively examined in correlation with fMRI (Hakun et al. 2008; Gamer et al. 2007, 2012). Insofar, some of the prospects for a scientifically valid forensic application expressed by competent researchers have a considerably more optimistic tone than the majority of voices dismissing the polygraph (Gamer et al. 2012, p. 513; Langleben et al. 2012; Langleben and Moriarty 2013; see also Vincent 2011). This is, however, vehemently contested terrain. Other scholars, by contrast, are no less concerned about fMRI applications in court than they are about polygraphy (Mobbs et al 2007). Some of them resolutely deny an admissibility of fMRI lie detection in court at least for the time being (Morse 2006, 2012; Greely and Illes 2007; Sinnott-Armstrong et al. 2008; Uttal 2009; Pustilnik 2009; Kanwisher 2009; Moriarty 2009; Brown and Murphy 2010; Oullier 2011; further references in Schauer 2010, pp. 1199–1202). Others even demand an outright regulatory ban on "any non-research use of new methods of lie detection, including specifically

fMRI-based lie detection, unless or until the method has been proven safe and effective to the satisfaction of a regulatory agency and has been vetted through the peer-reviewed scientific literature" (Greely and Illes op.cit. 413). Furthermore, to some extent, the controversy is reflected in a whole range of court decisions. Some of these have admitted fMRI in criminal proceedings (though mostly to no avail for the verdict), whereas (more) others have rejected it.[10] It appears impossible, at least for legal scholars, to take a fair, scientifically attestable and yet decisive stand on either one of the sides in this controversy.

There is, however, also no need to do so here in order to sensibly pursue the matter further. Most of the concerns voiced by skeptics relate to certain standards of scientific validity and reliability in expert testimony required for evidentiary purposes.[11] With regard to such standards, they find all currently available methods of fMRI for lie detection to be wanting in various respects. And from this perspective, a plausible prima facie case against fMRI lie detection can be made. Objections may invoke not only the technical difficulties of the scanning procedure or the intricate problems of developing effective test paradigms but also the complexity of the cognitive processes involved in the natural phenomenon to be investigated: human deceptive behavior.

Such behavior includes cognitive processes to generate intent and strategies of deception in a given context as well as executive processes to perform the chosen deceptive act (Luber et al. 2009). A useful, albeit still coarse, taxonomy distinguishes four categories of cognitive functions associated with deception: information management, risk management, impression management, and reputation management (Sip et al. 2007). Each of these activities itself encompasses a set of more elementary functions, such as planning, employing one's working memory, selecting between alternatives, and modulating response inhibition (Luber et al. 2007). All of these processes are assumed to interact in a systematic, though largely unconscious, way during deceptive behavior. Furthermore, in order to be accessible to fMRI investigations, this complex cognitive machinery must, at least roughly, be understood in terms of its neural underpinnings – a formidable task fraught with nescience and uncertainties. To cut this still extendable account somewhat short, if one simply assesses the potential of fMRI to disclose deception vis-à-vis such difficulties and against the backdrop of a universal standard of validity and reliability in natural sciences, the prospects for success certainly are rather bleak.

[10]Overviews in Jones and Shen 2012; Aronson 2010. Most noteworthy of these rulings perhaps the verdict in *Semrau* (US v. Semrau 2010; affirmed US Court of Appeals, 6th Cir. 2012; see also Gary Smith v. State of Maryland 2012). It is not only based on established juridical reasoning but also on an in-depth examination of the available scientific evidence from the perspective of evidentiary requirements in the criminal law.

[11]In science, "validity" and "reliability" are distinguished (though related); the former refers to whether research results really demonstrate what they purport to, whereas the latter to whether such results are consistently obtained in sufficiently equal experimental situations. In law, however, both terms are frequently used interchangeably. Nothing hinges on this terminology here; so in what follows, "validity" can usually also be read as "reliability," and vice versa.

There is, however, no such thing as a uniform, homogenous standard of scientific validity applying to all evidentiary purposes that are legitimately pursued by parties to criminal proceedings. This pertains especially to the three major players in court: the defendant, the prosecution, and the jurors or the judges. As is well known, the defendant, or his counsel, does not have to proof anything, neither his innocence nor truthfulness. Things are quite different from the perspective of prosecutors and, in the end, judges or juries. What stands in need of sound proof in criminal trials is every single precondition required for a guilty verdict: firstly, all relevant facts establishing that the deed in question was committed by the accused[12]; secondly, the absence of justifying circumstances, such as self-defense[13]; and finally, the preconditions of the defendant's personal culpability.[14] All of this must be established "beyond reasonable doubt" or (amounting to the same standard) "to the firm conviction of the court" by prosecutorial evidence. And none of its potential refutations falls on the part of the accused, not even to some weak standard of plausibility, let alone doubtlessness. On the other hand, he or she certainly has a right to proffer all sorts of possibly refuting evidence that might exonerate them. This decidedly asymmetric allocation of the burden of proof, a corollary of the fundamental principle of "*in dubio pro reo*" ("benefit of the doubt" for the suspect),[15] must have, one would surmise, some bearing on the question of standards of scientific validity that have to be met by the prosecution, on the one hand, and by the defendant, on the other, in their respective efforts to provide evidence for or against certain relevant circumstances.

Against this background, all a defendant needs to establish in order to evade conviction is some qualms in jurors' or judges' "firm" confidence in his guilt to the rather small degree that suffices to preclude the "beyond doubt" criterion. So he

[12]Which encompasses all objective facts constituting the commission of the offence ("actus reus") as well as the required subjective facts on the part of the defendant ("mens rea").

[13]Note that this absence of justifying circumstances is regularly *presumed* unless there are concrete indications to the contrary. In other words, the realization of the elements of crime of a particular offence, i.e., of a behavior that is *generally* forbidden on pain of punishment, indicates prima facie that it was also forbidden (unjustified) in the *concrete* case. So with regard to particular justifications, there is often no need to proof anything. But if there are indicia of the presence of justifying circumstances during the perpetration of the deed, the final burden of proof of their *absence* falls on the prosecutor.

[14]And, as the case may be, the absence of specific exculpatory circumstances. Not all legal orders contain provisions for such specific exculpations, but many do. For instance, according to Art. 17 of the German Criminal Code (GCC), an *inevitable* error on the defendant's part that his behavior is (or was) lawful exculpates him (but does not, of course, make the *deed* itself lawful); certain forms of non-justifying necessity (threats to life, bodily integrity, and personal freedom) also have this effect (Art. 35 GCC).

[15]Contained in various Human Rights Conventions in international law, e.g., Art. 6 para 2 of the European Convention of Human Rights (1950).

may well be interested in submitting evidence of his truthfulness whose scientific reliability appears rather shaky – if and when it is at least above a threshold of mere "junk science."[16] There is no doubt that at present the results of research in fMRI lie detection are already way above this threshold (even if, as yet, they have been derived almost entirely from basic research in laboratories and scarcely been applied to real-life situations). Thus, it can be fully appropriate for a defendant to proffer methods of expert testimony in favor of his veracity that the prosecution could not sincerely think of to proof the opposite. Such measures may well fit his own aims, while being entirely inept and hence inadmissible for prosecutorial purposes of determining guilt. The weak or contested scientific status of such methods of evidence would then be simply a matter of assessing their correspondingly weak probative value, but not a matter of their (in)admissibility.[17]

As to the polygraph, however, criminal courts in the United States as well as in Europe have tended to decide differently.[18] And with regard to the admissibility of fMRI, courts at least in the United States seem to more or less consistently stick to the principles of their reluctant judicature on polygraphy. That is to say, they appear to apply rather uniform standards of scientific reliability to expert testimony regardless of whether that evidence is obtained and assessed for purposes of convicting or of exonerating the accused. Hence, courts seem to ignore the fundamental difference in probative aims and burdens associated with prosecutorial purposes (and duties), on the one hand, and defense goals, on the other.[19]

Dangerous "Seductive Allures"?

Why this is so is not entirely clear. In its landmark decision *Daubert v. Merrell Dow Pharms, Inc.* (1993), the US Supreme Court recognizes that "scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes" (1993, at 591). This pertains to possible varieties of diverse *scientific* topics whose proof might be the aim of expert testimonies. But why not also apply this

[16]A criterion derived, e.g., from the US Federal Rule of Evidence 702; see *Best v. Lowe's Home Centers, Inc.*, (2009), at 176; a comparable threshold in Art. 244 para. 3 of the German Criminal Procedures Act ("entirely unfit and useless"). However, in accordance with the asymmetry principle in the evidence pointed out in the previous paragraph, the threshold of what counts as "junk science" also varies with regard to what evidentiary goal is at stake: either conviction (prosecutorial) or exonerating (defense-related) purposes of the evidence at hand.

[17]Of course, expert testimony introduced for either prosecutorial or defense purposes may ex post turn out to back up just the opposite purpose, respectively. Then, it must be taken by its probative value for what it objectively supports. Evidently, however, the question of admissibility must be decided ex ante: by assessing the probative value of the evidence for the aim it is intended to further by the party who proffers it.

[18]For the United States see Schauer 2010, p. 1196, n. 23; for Germany and, in passing, a few other European countries see Putzke et al. 2009.

[19]As to the jurisdiction in the United States, cf. the critical voices in the literature cited above; as to Germany, see, for instance, BGHSt 44, 308 (1998).

uncontested rule to the diverse *evidentiary* aims that the parties of a trial pursue? In both these situations, what is at stake are divergent goals which are supposed to be supported by the evidence proffered respectively. If it is true that "the same evidence may be extremely probative for one purpose and not even relevant for another,"[20] why not also concede that it may well be of a (albeit small) probative value for the purposes of one party and not even relevant for those of another?[21] Granted that the state of the art in fMRI for lie detection is "not anywhere near meeting the *Daubert* standard" (Seaman 2009, p. 933). It does not follow, however, that its use in court, if exacted by the appropriate party for their particular goals, is excluded. This not only accords with our considered judgments but also is corroborated by our clear intuitions. Consider the infamous decision of an Indian court that sentenced a woman to life imprisonment for murder in 2008. The verdict was crucially based on evidence from BEOS (brain electrical oscillations signature), a form of EEG brain scanning. It rightly strikes most researchers as well as most legal scholars as outrageous (cf. Giridharadas 2008; Aggarwal 2009). Now, by contrast, consider an acquittal of a suspect based on neuroimaging, for instance, the (hypothetical) ex post acquittal of the woman who, in a recent UK case, was convicted of poisoning a child in her care.[22] She served a prison term of several years but continued to assert her innocence. After her discharge, she was examined in a series of fMRI scans while being confronted with her own and, on the other hand, with the public prosecutor's version of the indicted incident. The results were markedly consistent with her own story. Thus, the scans, while certainly not "proving" that she was innocent, notably demonstrated that she may have been (cf. Spence et al. 2008). Few would find it scandalous had the case been retried on this new evidentiary basis and had the defendant then been acquitted on the "benefit of the doubt" principle.

Still, the postulate of differentiating the applicable standards of validity with regard to who demands an fMRI for what purpose in the trial does not meet with much approval by courts and many scholars. The reason for this seems to be something that is captured, for instance, in the US Federal Rule of Evidence 403. It permits courts "to exclude even relevant evidence if its probative value is substantially outweighed by a danger of confusing the issues or misleading the jury" (cf. US v. Semrau 2012, p. 17). Such a confusing and misleading potential is seen in the frequently invoked "seductive allure" that neuroimaging is believed to unfold vis-à-vis laypersons such as jurors in criminal trials (Weisberg et al. 2008). It has been termed "Christmas tree phenomenon" (Mobbs et al. 2007), alluding to the glittering colors of the computer-generated fMRI pictures indicating the differential blood flow in the examined brain areas. The idea seems to be that jurors or even judges might somehow be overwhelmed by such suggestive effects in their

[20]Brown and Murphy (2010, p. 1155), endorsing the cited evidentiary rule in *Daubert*.

[21]Cf. Merkel (2011, 244 pp.), see also the excellent exposition in Schauer (2010).

[22]The case was not retried, so there has not been an actual ex post acquittal of the convict, but it is not at all far-fetched that there could have been one.

ability to soberly assess fMRI evidence; hence, the admission of fMRI evidence would fly in the face of the Federal Rule of Evidence 403 (Sinnott-Armstrong et al. 2008).

Such concerns have an initial plausibility.[23] On second thoughts, however, they appear rather puzzling. If fMRI can provide relevant evidence for defense purposes, its admissibility is prima facie backed by a fundamental, in some jurisdictions constitutional, right of the defendant. This is obviously a matter of profound legal importance. If, on the other hand, there is a risk of seductive "Christmas tree" phenomena beclouding jurors' and even judges' capacities to soberly assess what's presented to them, why not then instruct jurors and judges how to avoid that risk and interpret the images properly (Feigenson 2006)? That is, why not explain the usually very limited probative value of such evidence even for the modest purposes of the defendant? Why preclude the entire option and thus run the risk of curtailing a defendant's basic rights?

One may suspect that another, rather clandestine intuition is at work here: one of fairness, stipulating a principle of "equality of weapons" between parties to a criminal trial. If the prosecution, so the idea might go, has no chance to introduce fMRI for purposes of conviction, then defendants should not be allowed either to avail themselves of such an opportunity for the opposite aim. But this again seems to rest on a misunderstanding of the roles of the parties in a criminal trial and the asymmetric tasks associated with these roles. From the perspective of the state and thus the public prosecutor, a criminal trial has (or, in any case, should have) nothing to do with a kind of adversarial contest which the prosecution should be determined to "win" by getting the accused sentenced and which therefore should be played on equal terms and premises.[24] Rather, a criminal trial is a procedure to publicly defend the validity of a broken social norm by making the person who is responsible for the breach "pay" for what they did, in order to symbolically reinstall ("repair") the norm's violated normative claim to universal obedience. Against this backdrop, the above-sketched asymmetry in evidentiary requirements between prosecutors and defendants gains its meaning as an imperative of justice, designed to protect notably the innocent, but also the culprit, in their legitimate interest to defend themselves on grounds of the principle *in dubio pro reo*. Hence, it should not be blurred by assimilating the standards of scientific validity required from both parties for the evidence they proffer, respectively.

[23]Notwithstanding their somewhat condescending attitude toward cognitive capacities of jurors, let alone judges, with regard to certain scientific niceties that all of the critics seem to understand without difficulty

[24]Notwithstanding the fact that most common law legal systems incorporate cross-examinations to utilize the dynamics of an adversarial interaction between prosecution and counsel in order to scrutinize the evidentiary material. This adversarial aspect has a purely auxiliary (instrumental) sense; it is not an expression of the genuine character of the trial – not more so, at any rate, than former (now fortunately discharged) methods of ascertaining matters of fact, such as torture. Its purpose is to effectively serve the primary goals of the trial as indicated above. It is not itself one of these goals but a method to achieve them.

Summing all this up, from the perspective of the prosecution and for the aim of conviction, expert testimony based on fMRI lie detection is useless and thus inadmissible. From the perspective of the defense counsel, however, the very same method of evidence may well be suitable and admissible. "Slight support (or weak evidence)," as Schauer puts it, "ought not to be good enough for scientists, but it is often sufficient for the law" (Schauer 2010, p. 1208). And here all the more so, since informal, clandestine, unprofessional, unreliable, and uncontrollable methods of assessing a defendant's or a witness' credibility in their testimony are common practice and entirely unavoidable in criminal trials. They include judgments on factors such as "whether a witness looks up or down, fidgets, speaks slowly or quickly, and speaks with apparent confidence" (Schauer 2010, p. 1213), or whether their voices vibrate, they slightly blush or pale, their foreheads show traces of perspiration or their eyelids flicker, etc. The intuitive and entirely unchecked ways of laypersons to interpret such signs as evidentiary clues of veracity or mendacity are certainly much farther off any objective standard of scientific validity than results of fMRI on lie detection. Yet no one could seriously consider their flat-out ban from courtrooms. In short, the admissibility of new methods of evidence on grounds of reliability cannot be assessed entirely independent of what has been admissible from time immemorial.

So the following prognosis appears plausible: Considering the present dynamics of fMRI research on lie detection, its methods will not, and should not, be banned entirely from courtrooms simply for reasons of their alleged infeasibility. Rather, they should be cautiously admitted on a case-by-case basis if proffered by defendants. This proposal, however, stands in need of a range of restrictions and controls, one of them personal, the others factual in character. They elucidate what "cautiously" in the foregoing remark means and requires. I will list them in the following subchapter. Note that we are still concerned here with questions of suitability and admissibility, not yet with problems of a normative in-principle justification of (or objections to) such methods.

Restrictions and Caveats: Personal and Factual

Here's the personal constraint. As yet, fMRI lie detection appears acceptable only on individuals that have freely consented to its use. None of the known fMRI procedures works to any satisfactory degree if applied against the will of subjects. If they oppose the procedure, they can disrupt its effects by covert countermeasures, i.e., behaviors deployed in order to defy deception detection. Possible countermeasures include physical means, such as biting one's tongue, or (primarily) mental strategies, such as recalling dramatic and arousing events in one's life or intensely deflecting one's attention onto other themes (cf. Ganis et al. 2011; Rosenfeld et al. 2004, on so-called P300-based imagings). With regard to defendants, this simply adds another argument to the one above against an application destined to obtain evidence for their conviction. And as to witnesses, as long as they can avail themselves of such simple ways of making the procedure entirely unreliable,

forced-on fMRI is an inadmissible evidence merely for this reason alone, apart from questions of the legitimacy of such force. However, new methods may arise in the foreseeable future that, with a sufficient degree of reliability, foreclose or restrict subjects' ability to even form mendacious thoughts (see Luber et al. 2009). Such methods would perhaps be apt to even forcibly obtain truthful information from witnesses and would, of course, at the same time pose serious problems of justifiability.[25] At present, however, their practical suitability for the purported aim must be denied.

Now let us take a look at the most obvious and most important factual *caveats*. Some of them have already been touched upon in the arguments of the fMRI skeptics quoted above.

1. Laypeople might, as critics fear, confuse fMRI pictures with something like direct snapshots of the brain during its engagement in cognitive tasks. This is false in various respects, as our above sketch of the scientific principles of fMRI already indicates. Such imagings are computer generated out of thousands of recordings of radio pulses, thus averaging their results over large numbers of single "resonances" of hydrogen protons during extended periods of time and over a range of broadly varying resonant signals. They do not directly show any neural activity (let alone at anyone point in time) but rather make it deducible in a twofold indirect manner: from radio echoes which in turn point to biomarkers that are not part of neural tissue proper but primarily reflect changes in metabolic processes in blood vessels. From these biomarkers, some more steps on the rather extended inferential route from fMRI scans to their courtroom applications still need to be taken. All of these steps require highly skilled professionals in the context of highly sophisticated experimental settings in order to lead to useful insights.
2. Furthermore, up to now by far the most fMRI data obtained in experiments and indicating the presence of either (relatively) normal or abnormal neural processes have their origins in the brains of numerous people, i.e., have been derived from, and averaged over, more or less large numbers of individual subjects in each experiment. This is also true for scans that show brains of people during lying as opposed to others speaking truthfully. The method is designed to "cross out" individual differences that appear on a rather broad scale of variations *within* groups of normal (in our case, truthful) or abnormal (deceptive) people as well as between such groups (Hariri 2009). This makes it difficult, if not impossible, to say with reasonable certitude of any one individual brain scan, which differs in some respects from (loosely speaking) the average "truthful" brain and resembles in just these respects the average "lying" brain, that it actually belongs to the latter group – and thus assures that its owner lied (Faigman 2010; Jones and Shen 2012, p. 356).[26]

[25]Hence, we will briefly return to this topic when discussing in-principle objections to fMRI for lie detection in court (see infra, section "Neuroimaging of Deception (2): Principled Normative Objections?").

[26]To avoid misunderstandings, brains do not lie; people do. The above wording serves as a shortcut only.

3. However, this difficulty of subsuming the results of an individual fMRI scan under a certain type of brain function, and hence of classifying the individual person as belonging to a certain group whose members (say, liars) show, on average, just that type of function, is not the only "translational" problem posed by fMRI. Two more are fairly obvious. One pertains to the highly stylized artificiality of the usually primitive lying tests in the scanner that are actually far off any halfway complex situation of deception in everyday life. And the other, related but more profound, pertains to the fact that people in the test situation are *requested* to lie by the researcher. Hence, their "lies" – if one wants to call such linguistic deviations from what's objectively correct "lies"[27] – are entirely free of stress for those who produce them and, hence, psychologically speaking, do not seem to bear much resemblance to the typical real-world lies (cf. Greely and Illes 2007, pp. 403–404; Editorial Nature Neuroscience 2008). What then do the results of such fMRI "lie test" studies indicate with regard to real-life situations in which the falsehood of a testimony, be it given mendaciously or negligently, may have grave consequences for the life of the individual who utters it? As of yet, no one really knows.
4. When people lie, numerous cortical areas, distributed widely over the whole brain, are involved. Typically, each of these areas is also involved in quite a few other mental activities, possibly very different in character than deceiving. There is no such thing as a "lying area" in the brain.
5. Even the conceptual contours of what is being searched for when a lie-detecting fMRI is applied are not clear. What exactly does "lying" imply? Saying "no" where "yes" would be appropriate? Just telling a somewhat different story from the true one? Omitting something? Concealing, slanting, shading, bending some of what is said? And a fortiori unknown is what kinds of differences in neural activity might correspond to such variations.[28] This difficulty may be overcome to a substantial degree, however, by employing (if suitable for the case in question) "concealed information tests" that do not require any verbal answer on the part of the subject, but only an automatic ("autonomic") response to certain visual stimuli on the part of his brain.

This is a formidable list of possible objections to the use of fMRI for lie detection in criminal trials. However, as we saw above, concerns about validity and reliability of the method can be overcome by first making the requisite distinctions and then, of course, taking all those possible objections into account. With regard to its scientific suitability only, no outright ban is called for.

[27]Which is denied, e.g., by Kanwisher (2009, p. 12). The conceptual question is not very interesting here (and can be decided either way); what is important, however, are possibly profound psychological differences between "lying" in the experimental setting and lying in an important real-world situation.

[28]One must not overlook, however, that these variations on truth and lying are "an omnipresent feature of modern litigation" (Schauer 2010, p. 1194) and thus must be dealt with in court in any case and can only be dealt with under considerable uncertainty, be it with or without the assistance (or distraction) of fMRI.

Neuroimaging of Deception (2): Principled Normative Objections?

There is, however, a further weighty concern, one from a normative perspective. Let's assume that someday fMRI for lie detection will be capable of ascertaining thoughts even from an uncooperative suspect with a high degree of certainty and hence well beyond any of the doubts about its validity discussed above.[29] Would it then be acceptable as evidence? Or would it violate the basic rights of defendants?

Again, we need to draw at least two distinctions in order to tackle this problem: one is between applications on defendants and on witnesses and the other between a compelled use and one consented to by the person concerned.

Compelled Application on Defendants?

Suspects have no legal obligation to actively contribute anything to their conviction. Consequently, they have a right to silence, which includes that no potentially incriminating testimonial evidence be obtained from them against their will or even by force. This privilege against compelled self-incrimination – in its classical Latin form, *nemo tenetur se ipsum accusare* ("no one is obligated to accuse himself") – is a fundamental principle of the law of criminal proceedings. It is also a positive legal norm in many jurisdictions and established in various international covenants.[30] If it comes to precisely determining its scope and content, however, notorious difficulties and long-standing doctrinal controversies arise.[31] They originate mainly from a potential normative conflict with another principle, an equally uncontested norm of procedural criminal law in most jurisdictions: Whereas a suspect must not be forced to actively testify against himself (*nemo tenetur*), his body in all its biological parts may well be examined against his will (if necessary forcibly) and thus be used as physical evidence against him. Put another way, whereas defendants cannot be compelled to reveal anything that might incriminate them with regard to the crime they are suspected of, every piece of their (inner and outer) body may be

[29]This is not sheer science fiction. It may become feasible in the not too distant future by first employing certain measures of brain stimulation that significantly impede a suspect's ability to engage brain networks involved in conscious deceit and then subsequently submit their brains to an fMRI via CIT, i.e., by scanning them for concealed knowledge or memory; cf. Luber et al. (2009).

[30]In the US Constitution, the principle is protected in the Fifth Amendment. The International Covenant on Civil and Political Rights (1966) expressly warrants it in Art. 14 para. 3 (g). And it is usually derived from Art. 6 para. 2 of the European Convention on Human Rights of 1950 ("presumption of innocence until proven guilty") – not a logical deduction, to be sure, but a plausible normative derivation.

[31]The arguments in these controversies are probably similar in most jurisdictions where the principle is guaranteed; they certainly are, for instance, in the US and the German scholarly debate; of the vast doctrinal literature in both these countries, cf. only Pardo (2008), Fox (2009), Verrel (2001), Eidam (2007) – each with a long list of further sources of reference.

forcibly scrutinized to reveal whatever may indicate their involvement in that crime. Obviously, these two principles may collide with each other time and again. Thus, they pose the problem of how to demarcate their respective normative purviews.

In decades of jurisprudential debate about this problem, a variety of differentiating criteria have been developed by courts and scholars, but none has proven entirely convincing or capable of achieving unanimous consent. In the US judicature (as an example for a common-law jurisdiction), the distinction between what is forbidden by the Fifth Amendment and what is not hinges on whether the compelled incriminating evidence from a suspect is "testimonial" or "physical," the former being unconstitutional whereas the latter doubtlessly legitimate (cf. Schmerber v. California 1966).[32] The respective distinction in the judicature of the German Federal Criminal Court (as exemplary for continental or "civil law" systems) rather turns on whether the suspect is compelled to *actively* contribute incriminating evidence of whatever kind and amount (forbidden) or only forced to *passively* endure bodily examinations of whatever kind or scale (allowed).[33]

Against the backdrop of these judicial criteria – and, in fact, of most other criteria proposed in scholarly discourse – the question whether fMRI scans for lie detection would violate the right to silence (*nemo tenetur*, Fifth Amendment) raises a puzzling problem. Is (forcibly) scrutinizing a suspect's brain in order to detect (via a series of inferences) certain mental entities or processes in their consciousness "physical" or "testimonial" in the sense deployed by US courts? That is, does it amount to nothing more than attaining information about physiological processes in one of their organs (their brain) – which would clearly be allowed and hence legitimate ground for whatever (perhaps incriminatory) conclusion to be drawn from it? Or is it rather akin to making them disclose (via their brains) knowledge, thoughts, and memories from their innermost mental sphere – which would just as clearly be forbidden by the *nemo tenetur* principle, or the Fifth Amendment for that matter? Or from the perspective of the German law, is what is extorted from the suspect an *active* contribution to their potential conviction (namely, cascades of brain activities) – which would be forbidden? Or are they only compelled to passively bear an examination of one of their physical organs (from which observers may then draw their own conclusions) – which would clearly be allowed?

In a sense fMRI for lie detection is simply both, depending on how one looks at it. It promises "distinctly testimonial-like information about a person's mind that is packaged in demonstrably physical-like form" (Fox 2009, p. 792). Epistemologically, such perspectivism, yielding different descriptions for the same object, is entirely unproblematic. Normatively speaking, however, fMRI cannot be both allowed and forbidden at the same time. So we must decide what we take it for.

[32]For a thorough examination of the "testimonial/physical" distinction and convincing critique of its shortcomings (see Fox 2009; see also Pardo 2008).

[33]It goes without saying that such an examination must not significantly threaten the suspect's health. For a largely convincing critique of this "active vs. passive" criterion see Verrel (2001).

This is a genuinely normative problem. It cannot be resolved by comparing phenomenal similarities and dissimilarities between fMRI and paradigm types of testimonial as opposed to physical evidence (or of actively contributing vs. passively bearing on the part of the suspect). Instead, we must clarify what deeper principle a suspect's right to silence is based on and then ask whether or not a compelled fMRI for lie detection would contravene that principle's basic normative sense. Simply, albeit correctly, ascertaining that compelled testimonial evidence (or compelled active self-incrimination) is unlawful in criminal proceedings does not explain why this is and should continue to be so and hence does not help us solve our problem.

We cannot enter here the labyrinth of intricate arguments that have been put forward to cope with this problem.[34] Suffice it to say the following: What decisively matters for the prohibition of forcing suspects to testify (actively and/or "testimonially") against themselves is the fact that such compulsion seizes the authority of control over thoughts, memories, knowledge, and other mental processes – in short, the inventory of one's mind and thus over the core of one's personality.[35] Exerting external control over someone's mind in such a way, be it by compulsive threats or by irresistible physical force, deprives a person of a constitutive element of personhood at large. That is to say, being a person in the full normative sense of the concept does not merely involve having a mental inventory of thoughts, reminiscences, emotions, etc. (all of which certainly constitute the individual "I") but also being in immediate command of the processes that dispose of and deal with such elements of the inner self (to the extent they are at one's willful disposal at all).

This is an evaluative, not a descriptive statement, and thus not accessible to scientific proof or refutation. It does, however, plausibly fit our concept of personhood. And it is also significantly confirmed by the historic background against which it should be judged: the so-called "inquisitional" type of criminal trial which for centuries was characterized by outright barbarous procedures of coercing suspects to confess, viz., by subjecting them to torture. The principle of *nemo tenetur* is designed to ban from criminal trials not only such methods but also the related goal of seizing control over a suspect's mind by gaining compelled access to their thoughts and thus depersonalizing them in a certain way and for a certain purpose. Hence, the innermost control over one's mind is declared legally sacrosanct from unwanted access and use by the state in criminal proceedings.

This provides us with a clear answer to our question: Compulsory fMRI for lie detection in suspects is illegitimate and excluded from criminal trials. For in such cases, the brain would not be searched to ascertain physiological activities, which in themselves are uninteresting for any evidentiary purpose, but to gain access to the corresponding mental processes by displacing the defendant as the subject of

[34]For exemplary scholarly analyses, see sources cited supra, n. 31.

[35]A very similar conception in Fox (2009, 796 pp.); relatedly for the German law Verrel (2001, 246 pp.).

control over these processes and their public proclamation. It is of no significance whether this happens via compulsion or via circumventing the controlling position of the individual altogether, i.e., by usurping their privileged access to their own mental sphere and thus extracting part of its content, as it were done in compulsory fMRI. This differentiation between neural activities and correlated mental processes does not rely on a philosophically flawed strong (or "substance") dualism of a Cartesian provenance, as Fox worries (Fox 2009, 793pp.). All monistic conceptions of the mind-brain relationship also presuppose a certain correlation (not interaction!) between mental states and their neural underpinnings, be it one of possible "reduction" of mind to brain processes, one of different "aspects" of one and the same entity, or any of the other conceptions proposed to grasp this specific correlational setting.[36] In our context, we may safely ignore these contested metaphysical problems, anyway. For our distinction between mental processes and brain activities is a purely normative one. Its adequacy does not depend on any particular philosophical position on the mind-brain problem.

Compelled Application on Witnesses?

Witnesses, of course, do not regularly have a right to refuse testimony (specific exceptions aside), much less a right to lie while giving evidence. Would it be legitimate to subject them to fMRI lie detection against their will? As long as they can evade the goal of the procedure by employing simple and effective countermeasures, its application is useless and thus inadmissible.[37] Scientific progress may, however, develop fMRI methods – perhaps, as the case may be, in conjunction with certain forms of brain stimulation (Luber et al 2009) – which are largely immune to such countermeasures. Then? There is no principled or (depending on the jurisdiction in question) constitutional objection that would decisively rule out such an application per se. However, in liberal states or (in Rawlsian terms) in well-ordered societies, the methods of compulsion must, of course, be restricted. Such states should not resort to physical force in order to coerce witnesses into brain scanners. And with regard to our above conclusion that controlling the access to one's own mind is a constitutive function of personhood, it is decidedly preferable for liberal states to also avoid other (nonphysical) forms of compelling fMRI for lie detection on disinclined witnesses.

[36]With, perhaps, the exception of an "eliminative monism" that attempts to somehow dispose of the mental side of brain processes altogether – not a very attractive philosophical position. It is unpromising to deny that we do subjectively experience mental events such as phenomenal states (of "what it's like"), even if they are nothing but "the other side" of brain processes (which we do not subjectively experience).

[37]As we saw above (II.3, supra), if they consent to the fMRI and want to testify in favor of the defendant's innocence, the procedure is not entirely unfeasible and hence may well be considered admissible.

fMRI for Lie Detection on Freely Consenting Suspects or Witnesses Who Request It?

There are no principled objections against such applications in criminal proceedings. And there is no reason to worry about an undue "functionalization" of persons and hence perhaps a violation of their dignity[38] through an fMRI they consented to. Their dignity begins with their choice. As to questions of admissibility, which are a different matter, we may now refer to our above arguments for a differential solution with respect to the different roles and the respective onus of proof of the parties to a criminal trial (see section "Neuroimaging of Deception: Feasibility? Admissibility?").

Conclusions: fMRI for Lie Detection

With regard to the admissibility and legitimacy of fMRI for lie detection, four conclusions from our foregoing considerations suggest themselves:

1. There should not be an outright ban, not even a present-day moratorium, on fMRI-based veracity tests in criminal trials. For prosecutorial purposes of conviction, however, the method is entirely unsuited and hence inadmissible. Not so, by contrast, for the much more modest goal of exonerating the accused under the principle of *in dubio pro reo* ("benefit of the doubt"). That the probative value of fMRI imaging will approach anything near certitude (comparable to DNA tests) in the foreseeable future is extremely unlikely. On the other hand, it is way above any dubious hyperboles of "junk science." Rather, it is an objective of serious research, and it will, in all likelihood, not stagnate at its current level. That this level should make fMRI for lie detection in forensic settings entirely unfeasible, i.e., even for the purposes of defendants and their counsel, is unconvincing.
2. However, the efficiency of fMRI tests can as yet rather easily be undermined by destructive countermeasures of unwilling subjects. Hence, they are only feasible, and thus admissible, if applied on freely consenting defendants or witnesses.
3. On the part of defendants, this holds also for the basic normative reason of their right to silence (principle of *nemo tenetur*) as protected in various international covenants and in national constitutions such as the Fifth Amendment in the United States. This principle forbids extracting possibly incriminating testimonial knowledge from suspects against their will by circumventing their personal control over their own thoughts and memories. This principle does not apply to witnesses. However, because controlling the access to one's own mind is a constitutive element of personhood, witnesses should not be compelled to undergo fMRI for lie detection either. If, by contrast, suspects or witnesses

[38] As was erroneously done in an early decision of the German Federal Criminal Court with regard to polygraphy that was requested by the defendant; see BGHSt 5, 332 (1954).

consent to, or even require, an fMRI test in order to strengthen their credibility, no principled objection stands in the way of complying with their request.

4. If fMRI results are admitted, it must be pointed out to jurors or judges by qualified experts that the value of fMRI as circumstantial evidence to assess the truthfulness of a testimony is usually low to marginal. This holds even for the modest purpose of exonerating the defendant under the principle *in dubio pro reo*. The reasons for this limited probative value must be elucidated to the triers. In particular, the above-listed caveats (see section "Restrictions and Caveats: Personal and Factual" (1–5)) should be expounded in order to counter whatever "seductive allure" or "Christmas tree" effect fMRI results might potentially exert.

FMRI for "Neuroprediction": Assessing Future Dangerousness

The necessity to make forensic prognoses about the future dangerousness of criminal defendants may occur in (at least) two different contexts: for most common law jurisdictions in the sentencing phase of a trial and furthermore (and in "continental" systems *only*) in procedures of preventive detention, which include the coerced confinement of violent sexual offenders as it is established in the so-called sexual predator statutes in many US jurisdictions (cf. Nadelhoffer et al. 2012, 75pp.). Could fMRI be employed for such risk assessments on potentially dangerous people?

Here is the premise that any sensible answer must reckon with: What is at stake in both these types of forensic predictions is the option for the state to impose sanctions on someone for something they have not done (but are only feared to eventually do in the future). It is clear at once that such a practice is somewhat of a borderline case for any legitimate legal order committed to principles of justice.[39] Against this background, it seems cogent that the state is obliged to utilize and exhaust all available, scientifically acceptable methods to ascertain the prognosis of the future dangerousness of a delinquent, given that such a prognosis is indispensable once the question of dangerousness has seriously arisen in a case. At present, it is based on the expertise of (usually two) psychiatrists, a cognitive basis fraught with uncertainties and all too often error (cf. Ennis and Litwack 1974; Thornberry and Jacoby 1979; Monahan et al. 2001) – or even on nothing but juries' or judges' intuitions, a base that's patently unfit to master a task of such importance (cf. Reidy et al. 2013).

[39]This throws into sharp relief concerns of justice about the established practice in many common law jurisdictions to impose harsher penalties than are matched by the degree of guilt (or viciousness, as it were) realized by someone's crime because the perpetrators are believed to pose a future danger to society. One should not be *punished* for something one has not committed (and hence should also not suffer an *extra* penalty beyond what their crimes alone make them deserve) – though one may certainly be *kept detained* if one is rightly assessed to pose a substantial risk of seriously harming others in the future.

Might fMRI scans provide adjuvant support here? Currently, there are at least two mental (or, for that matter, neural) predispositions to re-offend for which fMRI results might yield a sufficiently validated diagnostic and hence prognostic basis: pedophilia and psychopathy. Rather strong empirical evidence indicates a significantly higher risk of recidivism in perpetrators with one (let alone both) of these two mental conditions.[40] It seems that both these conditions can already be assessed with a sufficient degree of reliability by fMRI (for pedophilia cf. Ponseti et al. 2012; Wiebking et al. 2012; for psychopathy Wahlund and Kristiansson 2009; Anderson and Kiehl 2013). In light of these findings, states are not only entitled, but do indeed have an obligation to include "neuropredictive" methods in the prognostic procedures underlying the imposition of measures of preventive detention. Given that the question of dangerousness cannot be avoided, given furthermore how much is at stake for the person concerned (indefinite confinement for something they haven't done), legal decision makers certainly "need to be equipped with the best possible predictions concerning future dangerousness" (Nadelhoffer et al. 2012, p. 76).

Clear as this is, it still raises a few qualms and calls for a few caveats. Not among these qualms, by the way, are concerns about possible violations of the "*nemo tenetur*" principle (in the United States, the Fifth Amendment). That principle only grants protection against self-incriminating testimony with regard to a crime one is suspected to have committed. It does not, however, preclude any effort to obtain as many clues as possible (by legally approved means, of course) concerning their future dangerousness.[41] Here are three caveats:

1. Obviously, fMRI-based predictions cannot replace the classical psychiatric methods of risk assessment. They can only provide an additional or, as the case may be, complementary source of cognition in order to ascertain the broadest base feasible for predictions of potential risks of criminal recidivism.
2. It must be clearly pointed out what fMRI scans can and what they cannot demonstrate. That someone has pedophilic inclinations does not necessarily mean that they will in fact encroach upon children's sexual integrity. According to recent empirical research, quite a few more men (and even women) than ever commit a pedophilic crime do in fact have such inclinations but are able to restrain themselves vis-à-vis the threat of legal punishment (Wurtele et al. 2013). Thus, knowing that someone is a pedophile does not include foreknowing that they will sexually assault children, even if they have already committed at least one such assault in the past (as is regularly presupposed in cases of preventive detention). The same argument holds, and perhaps even more so, for violent offenders who are diagnosed to be psychopaths.

[40]For pedophiles, (see Wilson et al. 2011; Hanson and Morton-Bourgon 2005); for psychopathic offenders (see Olver and Wong 2006; Rice and Harris 2013); for psychopathic sex offenders (see Porter et al. 2009).

[41]The rather unfortunate fact notwithstanding that many common-law jurisdictions confound grounds for punishing past deeds with grounds for predicting future dangerousness (cf. supra, n. 39).

3. On the other hand, the empirical research mentioned above indicates a significant higher risk of future sexual assaults in people with pedophilic appetence than with other sexual, e.g., homo- or heterosexual, orientations. Furthermore, by far not all who sexually attack children are genuine pedophiles; some simply take advantage of the defenselessness of children, without a specific or exclusive sexual drive toward their young victims. Hence, being able to demonstrate with a high degree of reliability that someone who committed a sexual assault on a child really is pedophilic, as fMRI scans are apparently capable of, actually provides an important element for a sufficient cognitive basis to assess their risk of recidivism. And again, the same also holds for the prediction of future dangerousness of psychopathic violent offenders.

This warrants the following prospect: fMRI (and other) brain scans are or, at any rate, will certainly be well suited to contribute valuable assistance to the difficult task of prognosticating criminal recidivism in certain types of sexual or violent offenders. Thus, they might help us clear up some of the dark spots in the practice of preventive detention. We must, however, be careful not to allow them to rather add one further spot to that record: an excessive trust and hence a kind of mechanistic application of their results, based on an overassessment of their capabilities and an underestimation of their limits. There is no such thing as a "criminal brain." There are, however, mental dispositions to act in certain ways that raise the risk of becoming criminal in their possessors. Like all dispositions to act, they have their proximate (though of course not their only) causal source in people's brains. To identify them there and to draw legally relevant conclusions with the necessary skeptical diligence – that is the future task of courts, psychiatrists, and legal scholars. Probably rather sooner than later, the rapidly developing methods of neuroimaging will provide indispensable support.

Cross-References

- ▶ Ethical Issues in the Neuroprediction of Addiction Risk and Treatment Response
- ▶ Neuroimaging Neuroethics: Introduction
- ▶ Real-Time Functional Magnetic Resonance Imaging–Brain-Computer Interfacing in the Assessment and Treatment of Psychopathy: Potential and Challenges

References

Aggarwal, N. K. (2009). Neuroimaging, culture, and forensic psychiatry. *Journal of the American Academy of Psychiatry and the Law, 37*(2), 239–44.

Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for brain fMRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), *Functional MRI* (pp. 369–380). New York: Springer.

Anderson, N. E., & Kiehl, K. (2013). Functional neuroimaging and psychopathy. In K. A. Kiehl & W. Sinnott-Armstrong (Eds.), *Handbook on psychopathy and law* (pp. 131–149). Oxford: Oxford University Press.

Aronson, J. D. (2010). The law's use of brain evidence. *Annual Review of Law and Social Science, 6*, 93–108.
Ben-Shakar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the guilty knowledge test: A meta-analytic review. *Journal of Applied Psychology, 88*(1), 131–151.
Best v. Lowe's Home Centers, Inc. (2009). 563 F.3d 171 (6th Cir.).
BGHSt = Entscheidungen des Bundesgerichtshofs in Strafsachen, amtliche Sammlung (Decisions of the Federal Criminal Court of Germany, official collection).
Bigler, S. E., Allen, M., & Stimac, G. K. (2012). MRI and functional MRI. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry* (pp. 27–40). Chichester: Wiley.
Brown, T., & Murphy, M. (2010). Through a scanner darkly: Functional neuroimaging as evidence of a defendant's past mental states. *Stanford Law Review, 62*, 1119–1208.
Daubert v. Merrill Dow Pharmaceuticals Inc. (1993). *509 U.S. 579*.
Editorial (2008). Deceiving the law. *Nature Neuroscience 11*/*11*, 1231.
Eggen, J. M., & Laury, E. J. (2012). Toward a neuroscience model of tort law: How functional neuroimaging will transform tort doctrine. *Columbia Science and Technology Law Review, 13*, 235–306.
Eidam, L. (2007). *Die strafprozessuale Belastungsfreiheit am Beginn des 21. Jahrhunderts*. Frankfurt/Berlin: Peter Lang.
Ennis, B. J., & Litwack, T. R. (1974). Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *California Law Review 62*, 693–752.
Faigman, D. L. (2010). Evidentiary incommensurability. A preliminary exploration of the problem of reasoning from general scientific data to individualized legal decision-making. *Brooklyn Law Review 75*, 1115–1136.
Feigenson, N. (2006). Brain imaging and courtroom evidence: On the admissibility and persuasiveness of fMRI. *International Journal of Law in Context, 2*(3), 233–255.
Fox, D. (2009). The right to silence as protecting mental control. *Akron Law Review, 42*, 763–801.
Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping, 28*, 1287–1301.
Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2012). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social Cognitive and Affective Neuroscience, 7*, 506–515.
Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage, 55*, 312–319.
Gary Smith v. State of Maryland. (2012). http://mdcourts.gov/opinions/coa/2011/10a11.pdf. Accessed 17 Dec 2013.
Giridharadas, A. (2008, September 15). India's novel use of brain scans in courts is debated. *New York Times*.
Granacher, R. C. (2012). Potential uses of neuroimaging in personal injury civil cases. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry* (pp. 201–213). Chichester: Wiley.
Greely, H. T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law & Medicine, 33*, 377–431.
Hakun, J. G., Seelig, D., Ruparel, K., Loughead, J. W., Busch, E., Gur, R. C., & Langleben, D. D. (2008). fMRI investigation of the cognitive structure of the concealed information test. *Neurocase, 14*, 59–67.
Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154–1163.
Hariri, A. R. (2009). The neurobiology of individual differences in complex behavioral traits. *Annual Review of Neuroscience, 32*, 225–247.

Jones, O., & Shen, F. (2012). Law and neuroscience in the United States. In T. M. Spranger (Ed.), *International neurolaw: A comparative analysis* (pp. 349–380). Berlin/Heidelberg: Springer.

Jones, O. D., Buckholtz, J. W., Schall, J. D., & Marois, R. (2009). Brain imaging for legal thinkers: A guide for the perplexed. *Stanford Technology Law Review, 5*.

Kanwisher, N. (2009). The use of fMRI in lie detection: What has been shown and what has not. In E. Bizzi, S. E. Hyman, M. E. Raichle, N. Kanwisher, E. A. Phelps, S. J. Morse, W. Sinnott-Armstrong, J. S. Rakoff, H. T. Greely, et al. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 7–13). Cambridge, MA: American Academy of Arts and Sciences.

Langleben, D. D., Willard, D. F. X., Moriarty, J. C. (2012). Brain Imaging of Deception. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry* (pp. 217–236). Chichester: Wiley.

Langleben, D. D., & Moriarty, J. C. (2013). Using brain imaging for lie detection: Where science, law, and policy collide. *Psychology, Public Policy, and Law, 19*(2), 222–234.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*, 869–878.

Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic Resonance Imaging, 22*, 1517–1531.

Luber, B., Kinnunen, L. H., Rakitin, B. C., Ellsasser, R., Stern, Y., & Lisanby, S. H. (2007). Facilitation of performance in a working memory task with rTMS stimulation of the precuneus: Frequency and time-dependent effects. *Brain Research, 1128*, 120–129.

Luber, B., Fisher, C., Appelbaum, P. S., Ploessner, M., & Lisanby, S. H. (2009). Non-invasive brain stimulation in the detection of deception: Scientific challenges and ethical consequences. *Behavioral Sciences and the Law, 27*, 191–208.

MacLaren, V. V. (2001). A qualitative review of the guilty knowledge test. *Journal of Applied Psychology, 86*(4), 674–683.

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery and Psychiatry, 75*, 6–12.

McLaughlin, B., & Bennett, K. (2013). Supervenience. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/supervenience/. Accessed 15 Dec 2013.

Merkel, R. (2008). *Willensfreiheit und rechtliche Schuld. Eine strafrechtsphilosophische Untersuchung*. Baden-Baden: Nomos.

Merkel, R. (2011). Schuld, Charakter und normative Ansprechbarkeit. In M. Heinrich & C. Jäger (Eds.), *Strafrecht als Scientia Universalis, Festschrift für Claus Roxin zum 80* (Geburtstag, Vol. I, pp. 737–761). Berlin: De Gruyter.

Mobbs, D., Lau, H. C., Jones, O. D., & Frith, C. D. (2007). Law, responsibility, and the brain. *PLOS Biology, 5*, 693–700.

Monahan, J., Steadman, H. J, Silver, E., Appelbaum, B. S., Clark Robbins, P., Mulvey, E. P., Roth, L. H., Grisso, T., & Banks, S. (2001). *Rethinking Risk Assessment: The MacArthur Study of Mental Disorder and Violence*. Oxford: Oxford University Press.

Moriarty, J. C. (2009). Visions of deception: Neuroimages and the search for truth. *Akron Law Review, 42*, 739–762.

Moriarty, J. C., Langleben, D. D., & Provenzale, J. M. (2013). Brain trauma, PET scans and forensic complexity. *Behavioral Sciences and the Law, 31*(6), 702–720.

Morse, S. J. (2006). Brain overclaim syndrome and criminal responsibility: A diagnostic note. *Ohio State Journal of Criminal Law, 3*, 397–412.

Morse, S. J. (2012). Neuroimaging evidence in law: A plea for modesty and relevance. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry* (pp. 341–357). Chichester: Wiley.

Nadelhoffer, T., Bibas, S., Grafton, S., Kiehl, K., Mansfield, A., Sinnott-Armstrong, W., & Gazzaniga, M. (2012). Neuroprediction, violence, and the law: Setting the stage. *Neuroethics, 5*, 67–99.

National Research Council of the National Academies. (2003). *The polygraph and lie detection*. Washington D.C.: The National Academies Press.

Olver, M. E., & Wong, S. C. P. (2006). Psychopathy, sexual deviance, and recidivism among sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 18*, 65–82.
Oullier, O. (2011). Clear up this fuzzy thinking on brain scans. *Nature, 483*, 7.
Pardo, M. S. (2008). Self-incrimination and the epistemology of testimony. *Cardozo Law Review, 30*, 1023–1046.
Ponseti, J., Granert, O., Jansen, O., Wolff, S., Beier, K., Neutze, J., Deuschl, G., Mehdorn, H., Siebner, H., & Bosinski, H. (2012). Assessment of pedophilia using hemodynamic brain response to sexual stimuli. *Archives of General Psychiatry, 69*, 187–194.
Porter, S., ten Brinke, L., & Wilson, K. (2009). Crime profiles and conditional release performance of psychopathic and non-psychopathic sexual offenders. *Legal and Criminological Psychology, 14*, 109–118.
Pustilnik, A. (2009). Violence on the brain: A critique of neuroscience in criminal law. *Wake Forest Law Review, 44*, 183–237.
Putzke, H., Scheinfeld, J., Klein, G., & Undeutsch, U. (2009). Polygraphische Untersuchungen im Strafprozess. *Zeitschrift für die gesamte Strafrechtswissenschaft, 121*, 607–644.
Raichle, M. E. (2009). An introduction to functional brain imaging in the context of lie detection. In E. Bizzi, S. E. Hyman, M. E. Raichle, N. Kanwisher, E. A. Phelps, S. J. Morse, W. Sinnott-Armstrong, J. S. Rakoff, H. T. Greely, et al. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 3–6). Cambridge, MA: American Academy of Arts and Sciences.
Reidy, T. J., Sorensen, J. R., & Cunningham, M. D. (2013). Probability of criminal acts of violence: A test of jury predictive accuracy. *Behavioral Sciences and the Law, 31*, 286–305.
Rice, M., & Harris, G. T. (2013). Psychopathy and violent recidivism. In K. A. Kiehl & W. Sinnott-Armstrong (Eds.), *Handbook on psychopathy and law* (pp. 231–249). New York: Oxford University Press.
Robinson, H. (2013). Dualism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/dualism/. Accessed 15 Dec 2013.
Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology, 41*, 205–219.
Schauer, F. (2010). Can bad science be good evidence? Neuroscience, lie detection, and beyond. *Cornell Law Review, 95*, 1191–1219.
Schmerber v. California. (1966). *384 U.S. 757*.
Seaman, J. (2009). Black Boxes: fMRI lie detection and the role of the jury. *Akron Law Review, 42*, 931–939.
Sinnott-Armstrong, W., Roskies, A., Brown, T., & Murphy, E. (2008). Brain images as legal evidence. *Episteme, 5*, 359–373.
Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. (2007). Detecting deception: The scope and limits. *Trends in Cognitive Science, 12*(2), 48–53.
Smart, J. J. C. (2013). Mind/brain identity theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/mind-identity/. Accessed 15 Dec 2013.
Spence, S., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). 'Munchausen's syndrome by proxy' or a 'miscarriage of justice'? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 23*, 309–314.
Taylor, K. (2012). *The brain supremacy*. Oxford: Oxford University Press.
Thornberry, T. B. & Jacoby, J. E. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago: University of Chicago Press.
United States v. Semrau, US Court of Appeals, 6th Cir. (2012). http://www.ca6.uscourts.gov/opinions.pdf/12a0312p-06.pdf. Accessed 17 Dec 2013.
Uttal, W. (2009). *Neuroscience in the courtroom. What every lawyer should know about the mind and the brain*. Tucson: Lawyers & Judges Publishing.
Verrel, T. (2001). *Die Selbstbelastungsfreiheit im Strafverfahren*. München: C.H. Beck.
Vincent, N. (2011). Neuroimaging and responsibility assessments. *Neuroethics, 4*, 35–49.

Wahlund, K., & Kristiansson, M. (2009). Aggression, psychopathy and brain imaging - Review and future recommendations. *International Journal of Law and Psychiatry, 32*, 266–271.

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*, 470–477.

Wiebking, C., Sartorius, A., Dressing, H., & Northoff, G. (2012). Pedophilia. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry* (pp. 99–111). Chichester: Wiley.

Wilson, R. J., Abracen, J., Looman, J., Picheca, J. E., & Ferguson, M. (2011). Pedophilia: An evaluation of diagnostic and risk prediction methods. *Sexual Abuse: A Journal of Research and Treatment, 23*(2), 260–274.

Wurtele, S. K., Simons, D. A., & Moreno, T. (2013). Sexual interest in children among an online sample of men and women: Prevalence and correlates. *Sexual Abuse: A Journal of Research and Treatment*. Online first: http://sax.sagepub.com/content/early/2013/11/11/1079063213503688.full.pdf+html (last visited 10 January 2014).