Effects of Brain Lesions on Moral Agency: Ethical Dilemmas in Investigating Moral Behavior

Markus Christen and Sabine Müller

Abstract Understanding how the "brain produces behavior" is a guiding idea in neuroscience. It is thus of no surprise that establishing an interrelation between brain pathology and antisocial behavior has a long history in brain research. However, interrelating the brain with moral agency—the ability to act in reference to right and wrong—is tricky with respect to therapy and rehabilitation of patients affected by brain lesions. In this contribution, we outline the complexity of the relationship between the brain and moral behavior, and we discuss ethical issues of the neuroscience of ethics and of its clinical consequences. First, we introduce a theory of moral agency and apply it to the issue of behavioral changes caused by brain lesions. Second, we present a typology of brain lesions both with respect to their cause, their temporal development, and the potential for neural plasticity allowing for rehabilitation. We exemplify this scheme with case studies and outline major knowledge gaps that are relevant for clinical practice. Third, we analyze ethical pitfalls when trying to understand the brain–morality relation. In this way, our contribution addresses both researchers in neuroscience of ethics and clinicians who treat patients affected by brain lesions to better understand the complex ethical questions, which are raised by research and therapy of brain lesion patients.

Keywords Brain injury · Brain lesion · Neurodegenerative diseases · Moral agency · Neuroscience of ethics · Neuroethics

Abbreviations

DBS	Deep brain stimulation
DLPFC	Dorsolateral prefrontal cortex

M. Christen (✉)
University of Zurich, University Research Priority Program Ethics, Zollikerstrasse 117, 8008 Zürich, Switzerland
e-mail: christen@ethik.uzh.ch

S. Müller
Charité—Universitätsmedizin Berlin, Forschungsbereich Mind and Brain, Charitéplatz 1, 10117 Berlin, Germany
e-mail: mueller.sabine@charite.de

Curr Topics Behav Neurosci (2015) 19: 159–188
DOI: 10.1007/7854_2014_342

Published Online: 15 August 2014

FTD	Frontotemporal dementia
NMDA	N-Methyl-D-aspartate
PET	Positron emission tomography
PFC	Prefrontal cortex
SPECT	Single-photon emission computed tomography
ToM	Theory of mind
VMPFC	Ventromedial prefrontal cortex

Contents

1 Introduction

Case 1 The accident changed everything. One moment of inattention, and (Madison) fell down the scaffold, resulting in a severe head trauma. A complicated surgery and weeks of rehabilitation followed, until Madison could be discharged from the hospital. But Madison was not the same person any more. Soon, her marriage dissolved and she was unable to continue her work. Madison underwent several neuropsychological assessments to settle her health insurance claims. One time, she arrived in tears, because her grandfather just died. The neuropsychologist was very sorry and offered to cancel the meeting—but then Madison giggled, saying that this was just a joke. She sat down in her shorts, although it was winter. Her mood changed every minute—from exorbitant joy to deep sadness. She confabulated and was sometimes verbally aggressive. She insulted the neuropsychologists as one of the many

incompetent physicians she had met so far, unable to help her; and the next minute, argued that everything was fine with her and that she needed no help. Somehow she managed her life—she initiated relationships, but the relationships never lasted long. Sometime later, the neuropsychologist tried to contact her again, but the trail grew cold. None of the social workers who tried to help her after her accident knew where she was. No officials had any clue whether she was still living in town.

Case 2 Doctor Tolliver was a popular pediatrician—until police caught him in the act of abusing a 7-year-old girl during a medical examination. The police knew about other reports on sexual abuse of girls aged 3–12 years by Tolliver, some of them were even filmed by Tolliver. The investigators also found child pornography on the desktop computer of Tolliver. During interrogation, Tolliver claimed that since a year ago, he sometimes had an unstoppable drive to touch girls during medical examinations. While in investigative custody, neurological problems emerged and a brain tumor was diagnosed. Tolliver was successfully operated on before his trial. During trial, the defense counsel argued that the tumor caused pedophilia and additionally deficits in impulse control and emotion recognition, which were responsible for Tolliver's abusive actions toward children. He outlined a temporal correlation between tumor growth and the documented incidences of child abuse. However, the judge did not agree with this argument, since the scientific literature did not demonstrate a sufficiently deterministic relation between this type and location of brain tumor and delinquent behavior. Furthermore, Tolliver demonstrated an excellent ability to plan and organize the abuse and performed well as a pediatrician. Tolliver was found guilty and sentenced to 8 years in prison.

Case 3 It was almost 10 years ago when Sten was diagnosed with Parkinson's disease. In the early phase, the symptoms were well controlled by medication—but the disease progressed and it became increasingly difficult to avoid dyskinesia phases. In line with disease progression, Sten became depressed and apathetic; and his wife took care of him. She reduced her employment substantially and finally became his nurse. One day, Sten's physician explained that he might be a candidate for deep brain stimulation (DBS) and that this therapy could help diminish the side effects of medication. An assessment—demonstrating that Sten did not have a history of psychiatric disorders besides his Parkinsonism-related depression—demonstrated that DBS indeed was a suitable therapy for him. Sten decided on this option, and the surgery went well. The result was amazing, in particular for Sten. He felt that he gained a new life—but his wife could barely recognize him. Now, Sten often wanted to go out without her and he came back late. When checking his credit card bill, Sten's wife realized that her husband frequently visited strip clubs. Confronted with this fact, Sten admitted that he regularly visited prostitutes; he felt he had to catch up with all the life he had missed in the preceding years. He also admitted that somehow the DBS device might influence his new desires—but he rejects any change to his stimulation settings.

These three cases—all of them inspired by real patients—outline the complex relationship between changes in the brain and aberrations in morally relevant

behaviors. This complexity is present on both sides of the brain–behavior relationship. Behavioral changes may result from sudden injuries of the brain, slowly progressing brain diseases, or therapies intended to counteract brain disorders. While some behavioral changes are reversible by neurosurgery, medication, rehabilitation, natural healing processes, or adequate social surroundings, others are irreversible.

For some aberrant behaviors of brain lesion patients, it is difficult to evaluate the moral component of behavioral changes objectively, since no consensus exists about the morality of certain behaviors within a given society, and least of all between different cultures. Rather, moral evaluations of different behaviors depend on a given cultural context, differ between subcultures, and undergo transformation processes. By way of example, slavery is nearly undisputedly considered immoral today, whereas visiting prostitutes is discussed more controversially.

How to treat individuals living with damaged brains that influence their behavior poses ethical questions. Behavioral changes in individuals with a frontal lesion may be stressful for families and caregivers who live with them—and the social services and financial benefits in most modern societies may be inadequate for these patients.

This reminder of the complexity of the relationship between the brain and moral behavior is an important caveat against overly straightforward causal explanations of immoral behavior. Such a simplification is exemplified by the iconic figure of Phineas Gage—the railroad construction foreman who suffered in 1848 from one of the most prominent traumatic brain injuries in history. While using an iron-tamping rod to pack explosive powder into a hole, the powder detonated and the rod penetrated Gage's left cheek, tore through his brain, and exited his skull. Gage survived this accident, but became according to the popular narration, a different person. In the words of Dr. Edward H. Williams, the physician who treated Gage's injuries: "He is fitful, irreverent, indulging at times in the grossest profanity (which was not previously his custom), manifesting but little deference for his fellows, impatient of restraint or advice when it conflicts with his desires (…). His mind was radically changed, so decidedly that his friends and acquaintances said he was 'no longer Gage'" (Harlow 1868). Although most accounts of Gage's life after 1848 are strange mixtures of slight fact, considerable fancy and downright fabrication (Macmillan 2000), his case became a widely used example of how brain and moral behavior are related—that the dysfunction of some parts of the brain, namely the right orbitofrontal or ventromedial prefrontal cortex, inevitably leads to major aberrations in moral behavior. The case of Phineas Gage is frequently mentioned in the introduction of papers that discuss the relationship between brain and moral behavior.

Seen from a historical perspective, this relationship between brain lesions and (anti)social behavior is a recurrent topic in brain research. In 1888, Leonore Welt, the first woman in Switzerland who was allowed to study medicine at the University of Geneva, published what today would be called a review paper on character change after frontal lesions (Welt 1888). She discussed 11 cases—among them a case of her own clinical practice and the crowbar case referring to Phineas Gage—where frontal lesions were associated with negative changes in moral behavior. She also discussed 47 other cases, where such lesions did not have such effects. Certainly, degree and localization of these injuries were much harder to describe when

neuroimaging was not yet available. Nevertheless, Welt urged for caution when deterministically associating brain lesions with character changes.

To what extent this diagnostic caution is present today can be questioned. Experimental and clinical studies demonstrating that focal lesions in the right frontal cortex lead to specific changes in moral and social behavior generate almost four times more citations when compared to studies that describe the complexity of behavioral changes and social adaptations after frontal lobe injuries (Christen and Regard 2012). This citation bias may indicate an ethical dilemma associated with the relationship between brain and morality, namely that research may promote a neurodeterministic view of moral agency that is not sufficiently supported by the current state of knowledge.

In outlining this ethical dilemma, we have to resolve difficult measurement problems when analyzing the possible effects of brain injuries on moral agency, which denote the ability of individuals to act in reference to right and wrong. On the side of behavior, standardized questionnaires and tests (e.g., Iowa Gambling task, moral dilemmas) are available, but they may not reflect sufficiently the behavior changes and their effects in real life. And on the side of the brain, although sophisticated imaging techniques are used today, it is still difficult to directly assess residual functionalities in the affected brain tissue, particularly in cases involving neurodegenerative diseases, and the potential of neuroplasticity.

We structure the investigation as follows:

1. We need a detailed understanding of moral agency. This involves both an empirical part—namely outlining mental competencies and the related physiological conditions—and a normative part. The latter is needed to evaluate the legitimacy of moral claims toward the behavior of others.
2. We need a typology of brain injuries and their known behavioral sequelae. A major issue is the variability of behavioral sequelae of brain lesions. This variability may be partly explained by difficulties of assessing the exact location of lesions and their effects on neural networks. Furthermore, the variability could be based on individual differences in neuronal plasticity and differences in rehabilitation measures as well as on differences in the pre-lesion personality and social relationships.
3. We have to keep in mind that the endeavor of relating brain and behavior itself has a history and may be driven by different agendas. Ethics research is not a purely objective or rational science. Ethical justifications appeal to intuitions that have both natural and cultural histories. Thus, the ethical framework used to investigate moral behavior itself needs to be reflected upon as well.

The structure of our contribution to this volume follows this basic outline: In Sect. 2, we introduce the notion of moral agency both with respect to its normative and empirical dimension. In Sect. 3, we provide a typology of effects of brain lesions on moral behavior. In Sect. 4, we discuss ethical pitfalls of relating brain lesions with moral behavior changes. Section 5 concludes our contribution with some preliminary thoughts on using knowledge on the relation between brain and morality to restore moral behavior that is compromised by a brain lesion.

2 Moral Agency and the Brain

Human beings possess the ability to act with reference to right and wrong, which is framed as moral agency. The structure of moral agency as we construe it here is threefold (Christen and Alfano 2013). First, moral agency requires a specified set of competencies that the agent must have. Second, it involves a normative reference frame to which the agent has at least partial access. Third, moral agency is always situated in a context that consists of other agents and physical boundary conditions that constrain behavior. Competencies, normative frame, and context thus form the structural components of moral agency. A particular empirical investigation of moral agency may refer to just one or two of these structural components or to the interaction of two or all three components.

2.1 Conceptual Issues of Moral Agency

A relevant problem refers to the prescription of agency. In theory, most would agree that behaviors caused by defined brain lesions or abnormalities are not under the control of the individual and thus cannot be called acts or considered as an expression of moral agency. But in practice, brain lesions or other abnormalities often do not prove a lack of understanding the wrongness of a certain act at the time of commission (Batts 2009). Neither does it prove an absent ability to act according to one's insight, which would justify a diminished or missing criminal responsibility in most European countries. Therefore, in most cases, it is not possible to draw a clear causal line from a brain lesion or other abnormality to a missing moral agency and thus to a missing moral and legal responsibility. Another controversial issue is to what extent patients with ventromedial frontal damage possess moral concepts or relevant moral beliefs (Cholbi 2006; Roskies 2006), a discussion that relates to the philosophical debate on internalism and externalism (for an overview see Björklund et al. 2012). These important issues refer to the broad discussion on free will and the determination of behavior by the brain. This problem can have practical consequences, as our second case with Tolliver outlines, where the counsel was unable to convince the judge that the defendant's brain tumor caused him to abuse children sexually. We will come back to this point in Sect. 4.

For the following explanations, we propose to relate moral agency to the fact that patterns of moral behavior are displayed by persons whose behavior is somehow regulated by a normative framework that includes an idea of good and bad. We use the term "patterns of moral behavior" rather than "moral actions" in this context because we do not want to restrict it to punctate actions. This is consistent with the clinical practice of therapy and rehabilitation that does not address specific acts but behavioral patterns and dispositions.

The first thing to specify is what the term "moral" should denote. A simple fact about morality is that people are disposed to react to issues according to what they

consider right or wrong, good or bad. This implies the existence of some normative frame and its connection with the real world in the sense that it guides thought, feeling, deliberation, and behavior of people. Another basic fact is that morality is situated in a social world of actions, judgments, negotiations, and other kinds of expressions made by social beings. This social world is embedded in a history, and its evolution is driven by many different factors. This means that acts, norms, and virtues that we may call moral are subject to fuzziness in two respects: First, within a society, there are actions that are undisputedly either moral or immoral, whereas other actions are less clear in that respect. Second, across societies and during history, the moral condemnation of some behaviors seems to be stable, whereas others undergo remarkable changes. Thus, moral evaluations of given actions differ both with respect to inner-societal agreement and evaluation stability over time.

Figure 1 illustrates these two dimensions with exemplary cases, although the precise location of these acts in this scheme can be debated. Morally, condemning the murder of innocent people is relatively stable both within a society as well as during the course of time. Slavery was for a long time morally accepted within societies but lost acceptance in a relatively short time span and is now regarded as absolutely inacceptable (Appiah 2010). Abortion has a long history of moral disagreement, whereas each position is relatively stable in time. Finally, the degree of moral acceptance of homosexuality shifted several times in history and to date the inner-societal disagreement is still high in many countries.

When evaluating the changes in moral behavior of patients, both the evaluation instability of moral behaviors and the inner-societal disagreement about them have to be taken into account. For the following general discussion, we define morality very broadly as a set of norms, principles, values, and virtues that are governed by an orientation toward the good. As such, morality reflects respect and concern for oneself and for other entities (persons, animals, or environment) and is embedded in a justification structure. We are aware that understanding one's moral decision-making

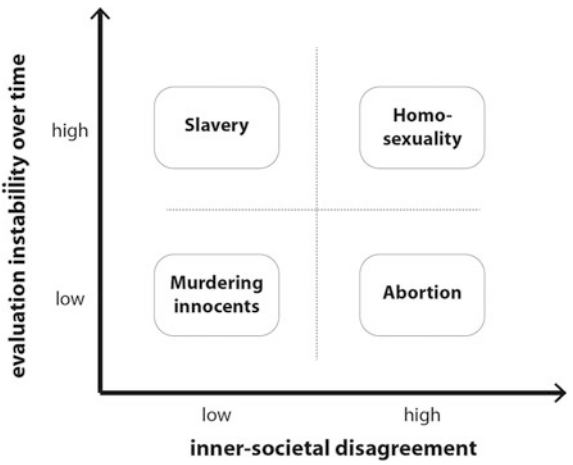

Fig. 1 Exemplary cases of moral valuation structured along the dimensions inner-societal disagreement and evaluation instability over time. The figure only identifies ideal cases in the four quadrants of the scheme, separated by a *dotted line*

and behavior requires an analysis of the agent's understanding of morality and on what he or she considers right or wrong. In addition, one would have to assess the actual justifications and their adequateness for an analysis of arguments.

2.2 Moral Intelligence as a Psychological Working Model

In our topic, a natural focus would be on the competencies and their foundation in the brain. This requires a theoretical framework that summarizes our knowledge on how agents reason, decide, and act morally. The major source of this knowledge is still moral psychology, which underwent a remarkable development in the last few years. In the following, we propose to use the concept of moral intelligence (Tanner and Christen 2013) as a theoretical framework.

Moral intelligence is defined as the capability to process moral information and to manage self-regulation in any way that desirable moral ends can be attained. It refers to the set of skills the moral agent needs in order to align her behavior with the moral ends she has set for herself, using the broad understanding of morality defined above. It is thus a skill-based conception of moral behavior, analogous to the concept of emotional intelligence that describes the ability to deal with emotions. The framework describes the sequential logic of moral behavior along with the associated underlying psychological processes, and the way in which implicit and explicit knowledge of morality and its justifications are included. These elements underlie the five competencies of moral intelligence (see also Fig. 2):

- *Moral compass*: This metaphor encompasses the set of moral schemata whose content is responsible for orienting the subject's behavior (Narvaez 2005). As such, it is concerned with mental representations of both declarative and procedural knowledge, each of which is accessible to the subject in varying degrees.
- *Moral commitment*: The ability to activate or sustain a motivation for the inclusion of moral considerations in the process of perception, decision-making, and action. In contrast to the typical process logic of moral behavior (perception → decision → motivation → action; Rest 1986), moral commitment is a

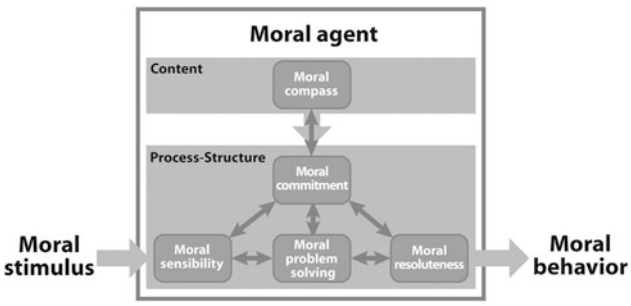

Fig. 2 The five building blocks representing competencies of moral intelligence in relation to the multistage model of moral functioning (adaptation from Tanner and Christen 2013)

capacity that influences all stages of the process, and in particular provides a motivational force to the semantic content of the moral compass.

- *Moral sensibility*: The ability to recognize morally salient aspects of a particular situation. The relevance of moral sensibility is obvious: If morally relevant aspects of a situation are not recognized, there is no cause to be concerned with the question of right action.
- *Moral problem solving*: The ability to bring the morally salient features of a situation to the decision-making process, and depending on the degree of conflict involved (e.g., if the problem has the structure of a dilemma), to arrive at a decision consistent with the subject's particular moral compass.
- *Moral resoluteness*: The ability to carry out one's own decisions despite, inter alia, external or internal resistance and barriers.

The concept of moral intelligence integrates the findings of (moral) psychological research into a unified model. As such, it enters an area with a rather long tradition. What distinguishes our model from other approaches is the central role of moral commitment, i.e., the capacity to uphold the demands of morality throughout this entire process and to align one's cognitions, decisions, and actions with one's moral ends. Moral commitment is to some extent the bridge between the moral compass and the other competencies of moral intelligence, and expresses the will to apply the contents of the moral compass.

It is unlikely that the building blocks of moral intelligence are related to distinct and clearly separable neuronal modules, because it is generally questionable whether mental processes can be defined and separated in a way that permits them to be associated with particular brain regions (Uttal 2001). The competencies also differ in their degree of exactness: Whereas moral sensibility is conceptually the simplest component, moral problem solving and moral resoluteness are more rich theoretical constructs. In addition, the psychological literature on the five competencies is not equally well developed. Moral problem solving or decision-making and, to a lesser degree, moral commitment within motivation psychology have been the subject of research for decades, in particular within developmental moral psychology advanced, among others, by Jean Piaget (1932) and Kohlberg (1981). Moral sensibility and moral resoluteness, however, are less well studied. Despite these difficulties, the framework of moral intelligence provides more precise considerations of which competencies may be affected by brain lesions, although one cannot expect that a specified lesion affects only one of those competencies, leaving the others intact.

2.3 Problems in Interrelating the Brain with Moral Agency

Our current knowledge on the neuroscience of ethics supports the expectation that there is no one-to-one correspondence between clearly discernible brain structures and functions on the one hand and the competencies that outline moral intelligence,

or even moral agency in general, on the other hand. The so-called moral brain obviously consists of a large functional network including both cortical and subcortical anatomical structures (recent overviews: Mendez 2009; Fumagalli and Priori 2012; Pascual et al. 2013). Because moral agency is based on a complex process, these brain structures share their neural circuits with those controlling other mental processes, such as emotions, motivations, decision-making in general, impulse control, and theory of mind (ToM). A moral brain does not exist per se; rather, many subsystems of the emotional and the cognitive brain systems are engaged in moral processes. These complex processes are influenced by many genetic, endocrine, and environmental factors (Fumagalli and Priori 2012; Pascual et al. 2013).

Among the anatomical structures implicated in moral agency are the frontal, temporal, and cingulate cortices; i.e., considerable parts of the cortical hemispheres: The prefrontal cortex (PFC) regulates activity in subcortical emotional centers and is involved in planning and supervising moral decisions. When its functionality is disturbed, impulsive aggression becomes more probable (Fumagalli and Priori 2012). Patients with prefrontal lesions—especially in the orbito-prefrontal and medial regions—are often significantly impaired in both cognitive and affective empathy (Shamay-Tsoory et al. 2004; Eslinger et al. 2004). Patients with bilateral lesions of the orbitofrontal cortex show impairments in social behavior (Hornak et al. 2003). Furthermore, changes in moral decision-making have been found in prefrontal lesion patients (Koenigs et al. 2007; Ciaramelli et al. 2007). Based on such studies, it has been claimed that the ventromedial prefrontal cortex (VMPFC) attaches moral and emotional value to social events and anticipates their future outcomes. It is involved in ToM and empathy, mediates automatic moral and prosocial reactions, and participates in social emotions, including guilt, embarrassment, and compassion. The temporal lobes are also involved in ToM, and their dysfunction is often implicated in violent psychopathy. The dorsolateral prefrontal cortex (DLPFC) is involved in cognitive empathy and in the application of reasoned analysis to moral situations (Mendez 2009). The cingulate cortex mediates conflicts, particularly conflicts between emotional and rational components of moral reasoning (Fumagalli and Priori 2012). Current research suggests that the cortical structures most directly involved in abnormal moral behavior are the right medial orbitofrontal cortex and the right ventromedial prefrontal cortex (Fumagalli and Priori 2012). Subcortical structures are also involved in moral behavior, particularly the amygdalae, the hippocampus and the basal ganglia (Mendez 2009; Fumagalli and Priori 2012; Pascual et al. 2013).

Although the relationship between frontal lobe damage and morally relevant behavior aberrations like aggression has been part of clinical experience for more than 60 years, most of the evidence is case-based. There are only a few larger studies with appropriate design (Hawkins and Trobst 2000). The most known retrospective study examined the relationship between frontal lobe lesions and aggression in 279 veterans who had sustained penetrating head injuries, which were compared to 57 veterans without brain injury matched by age, education, and time

served in Vietnam (Grafman et al. 1996). They found that the veterans with brain injury were more aggressive than control veterans, as reported by family and friends. In particular, veterans with ventromedial frontal lobe lesions were reported to be most aggressive, when compared to veterans with lesions elsewhere in the brain. But the authors also reported that "not all patients with these lesions had such behavior, and some patients with lesions elsewhere in the brain, and even normal controls, can show an increased tendency toward aggressive and violent behavior" (Grafman et al. 1996, p. 1237). There are also other case studies of patients with massive frontal lesions that are not compatible with a clear causal link between lesion and lasting behavior changes. For example, some patients do not show aberrant social behavior despite the lesions (Feinstein et al. 2010), in some patients the behavior changes after the lesions are reversible (Frías Ibáñez et al. 2008), and in some patients the behavioral and personality changes are compatible with stable functioning in family, professional, and social settings (Mataró et al. 2001). Therefore, the available evidence does not provide conclusive evidence that frontal lesions inevitably lead to such behavior changes. A too schematic, one-to-one connection between lesions in specific brain areas and specific moral behavior aberrations is misleading.

One reason for this variety found in the literature on the interrelation between the brain and moral behavior refers to the experiments that are used in these studies. Currently, a gross variety of tasks is used for assessing morality in the context of moral psychology or the neuroscience of ethics, which makes it difficult to compare the results of these studies. Furthermore, most moral tasks have intrinsic limitations. For example, they are not ecologically valid in that they reflect environmental and daily experience only poorly, or they request abstract judgments that exclude the complex decisional context. Additionally, task instructions usually forbid the subjects to make additional assumptions not included in the text, even though problem solving automatically intervenes in these situations. Finally, moral items distinctly differ from one another and involve different moral rules, violations, and values such as honesty, money, life, health, probity, or solidarity.

An important methodological limitation of most experimental studies is that they focus on moral judgments; i.e., the researchers account for, predict, or find neural correlates to moral judgments that they use in their experiments. These moral judgments are of a specific kind and have several defining features (Abend 2013). They are made in response to specific stimuli in imaginary situations, and they use only thin ethical concepts such as: okay, appropriate, permissible, acceptable, wrong, etc. In addition, they are fixed, verdict-like, and clear—not conceptually or semantically muddled, incoherent, etc. But moral judgments do not only occur as responses to specific stimuli or eliciting situations. Rather, some moral judgments develop over longer periods and are based on the reflection of many experiences and theoretical deliberation. In addition, morality cannot be reduced to moral judgments. This problem concerns in particular virtue ethics—an ethical approach that evaluates the character of persons in contrast to approaches that evaluate their actions, either in terms of duties or rules (deontology) or their consequences (consequentialism). Moral evaluations of actions are more easily expressed by

moral judgments. Abend (2013) argues correctly that the object of study of much recent work on the connection between the brain and morality is not morality per se, but a particular kind of individual moral judgment.

And even within this special sample, complexity remains. Parkinson et al. (2011) investigated moral scenarios that involved disgusting, harmful, and dishonest behavior along with a neutral scenario, and asked subjects to judge the general moral wrongness of the actions within each scenario as well as the degree of disgust, harm or dishonesty while in a fMRI scanner. They found that the latter three statements were subserved by distinct neural systems and these differences were much more robust than differences in wrongness judgments within a moral area. The dorsomedial prefrontal cortex was the only region activated by all scenarios judged to be morally wrong in comparison with neutral scenarios. However, this region was also activated by dishonest and harmful scenarios judged not to be morally wrong. Furthermore, these scenarios were not suggestive of a domain-general role that is neither specific for nor predictive of moral decisions. The results suggest that moral judgment is not a wholly unified faculty in the human brain, but rather, instantiated in dissociable neural systems that are engaged differentially depending on the type of transgression being judged.

In summary, this brief overview suggests that the attempt to find clear-cut connections between a fine-grained understanding of moral agency and defined neuronal structures may lead to a picture that is too complex to be useful in a clinical context. For example, there may be different neuronal systems that are responsible for moral sensibility related to harm versus moral sensibility related to honesty. A focal lesion may thus impair one aspect of moral sensibility more than others—but the relevance of this imbalance will depend on the situation in which this competence is needed. In addition to impairing one aspect of moral sensibility, a focal lesion may also influence other competencies (maybe also nonmoral ones), as the affected brain region serves many basic functions. This complexity, however, is usually not assessed in experimental studies involving lesion patients because it is not feasible to perform a full evaluation of all possible impairments a brain lesion may cause.

The basic problem (see Fig. 3) is that on the side of the phenomenology of moral behavior, one needs a sufficiently elaborated but not too complex set of constructs that describe competencies relevant for moral behavior, such as the moral intelligence model. A rehabilitation specialist can neither work with a too general concept such as moral behavior, nor with a too fine-grained understanding of moral behavior such as impairment of honesty-related moral resoluteness. On the side of the investigation of the lesions, a similar problem emerges: One needs a partitioning that is compatible with the size of regions that can be affected through focal lesions —and with the methods available to actually identify regions and their degree of impairment. It is no coincidence that the current neuroscience of ethics denotes still rather large regions as being relevant for moral behavior, for example, the orbitofrontal cortex, which extends over several square centimeters. Taking all methodological issues of properly identifying such regions aside, it is clear that they are involved in many basic functional networks that may be clearly identified sometime

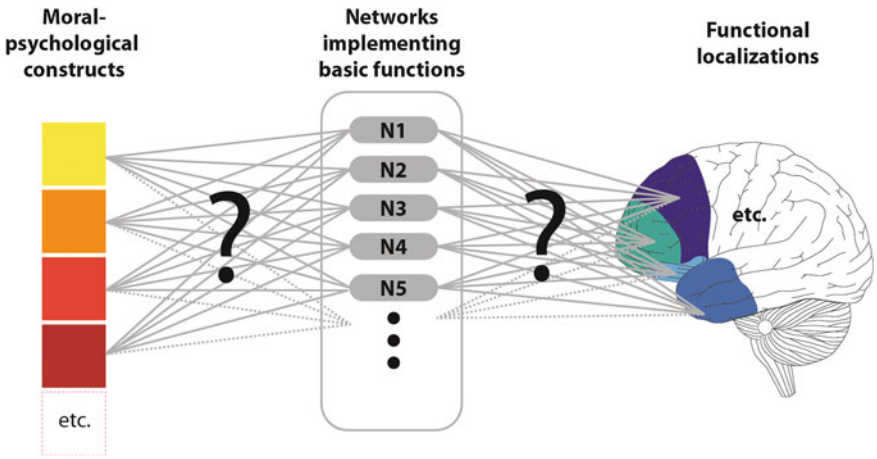

Fig. 3 Illustrating the problem of connecting moral–psychological constructs that describe relevant and usable moral competencies (*left side*) and functional localizations (*right side*) that actually involve many networks (N1, N2, N3, etc.) that may be affected by a lesion

in the future. This will be a challenging endeavor, as it is still rather unclear what constitutes a basic function and what should be the demarcation criterion within the huge cortical networks. Several of those networks that implement basic functions will then be recruited in order to form a defined moral–psychological construct that is useful, for example, in rehabilitation.

A way out of this problem is to resign from an elaborated phenomenology of moral agency and to focus on very few behavior types that seem to have clear moral impact, like violence. Fumagalli and Priori (2012) write: "From a behavioral point of view, the major consequence of moral abnormality is violence," which stands exemplarily for this position. However, as we will outline in Sect. 3.3, this position also raises ethical questions. We now proceed by providing a typology of brain lesions that may affect moral behavior.

3 Effects of Brain Injuries on Moral Behavior

Various types of pathological processes can affect the brain in a way that produces changes in behavior. Some of them occur instantly such as in trauma or stroke; others develop over a longer time scale, for example tumor growth or neurodegeneration. In the following, we use the notion of lesion or damage in a rather general way to indicate any kind of structural damage to brain tissue that have functional consequences. Examples of brain lesions include the following:

- Direct injury of brain tissue (e.g., gunshot)
- Ischemic damage to brain tissue (e.g., stroke, aneurysm rupture)

- Tumor-related damage to brain tissue (e.g., damage due to infiltrating tumor growth or expansion lesion due to increased pressure)
- Neurodegenerative processes (e.g., death of specific cell types as in Parkinsonism)
- Brain inflammation (e.g., encephalitis)

For analyzing differences between types of brain injuries, we classify them along two dimensions: the temporal scale of the brain injury and the plasticity potential of the brain injury.

The first dimension describes the typical temporal course of different types of brain injuries, namely the temporal course of their onset, of the subsequent development of changes in personality and behavior, and of the necessary therapies and rehabilitation processes. For example, the type of brain tumors determines the occurrence of symptoms (suddenly or gradually), the duration of necessary therapies (several hours for tumor resection or life-long for pharmacotherapy) and of necessary rehabilitation (short training course or life-long training). This dimension also influences how other people, particularly from the direct social surrounding of the patient, will react to lesion-related changes. For example, personality changes that develop slowly allow the family a better customization to changed behavior of the patient.

The second dimension is the plasticity of the brain that may allow for a partial or full reversibility or compensation of functional losses. This dimension comprises both healing processes of the affected brain tissue and functional shifts. An example for the latter is the transfer of the language centers from the left to the right hemisphere after resection of the left hemisphere due to Rasmussen encephalitis causing therapy-refractory epilepsy, which has been reported only from children younger than 5–6 years (Varadkar et al. 2014). Several factors influence the plasticity of the brain:

- The patient's age at the time of a brain lesion
- The exact location of the lesioned area and its physiological functions
- Healing processes
- Compensation processes (e.g., shift from certain functions to another than the affected area)
- The therapy and rehabilitation measures used (including medication) and their efficiency

Figure 4 provides an overview on these two dimensions. On the x-axis, brain injuries are sorted according to their temporal scale, that is, whether they develop fast or slowly. On the y-axis, brain injuries are sorted according to their plasticity potential. These two dimensions are relevant for brain damage resulting from pathological processes (white boxes) and from interventions as unintended side effects (gray boxes). Again, ideal types are shown, and the location of each example within each of the four quadrants is not intended to be precise.

Using this classification, we will now provide a case-based overview to outline the diversity of moral behavior changes due to brain lesions.

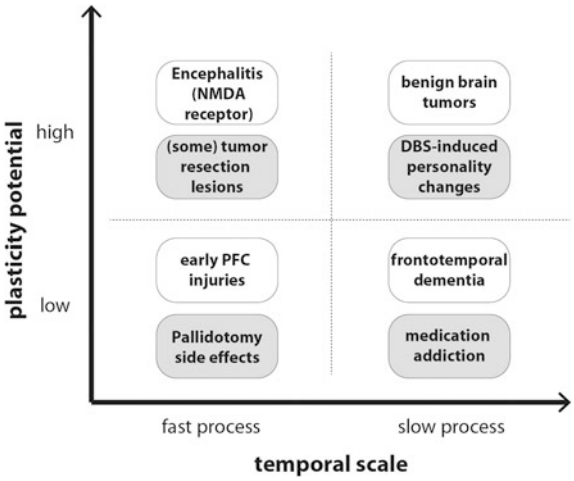

Fig. 4 Examples of pathological processes or interventions side effects influencing the brain. The examples are structured along the dimensions temporal scale of the process and the plasticity potential (potential of functional reversibility of the lesion sequelae). The *white boxes* denote changes due to pathological processes, the *gray boxes* non-intended changes due to interventions. The figure only identifies ideal types in the four quadrants of the scheme (*dotted line*)

3.1 Fast Processes with High Plasticity Potential

If changes in personality and behavior are caused by fast developing brain lesions which are reversible by adequate therapies or by natural healing processes, then for the patient and people in his/her social surrounding it becomes obvious that the changes were caused by a brain lesion, and not by the patient's "evil will." Such cases may be caused either by disease, injury, or therapy. They are particularly interesting since they allow us to study causal relationships between brain lesions and changes in personality and behavior in a bidirectional way.

An impressive example is NMDA receptor antibody encephalitis (NMDA: N-Methyl-D-aspartate), which was first described in 2007 as an autoimmune disease characterized by rapid development of psychosis, paranoia, aggressiveness, and other symptoms which may lead to a misdiagnosis of schizophrenia. Fortunately, the inflammation-caused mental sequelae are mostly reversible with timely administration of an effective therapy consisting of cortisol administration, hemodialysis, and immunotherapy (Dalmau et al. 2007, 2008). But since this disease has not been discovered before 2007, and since new scientific findings need some time for clinical translation, it is likely that many patients suffering from psychosis have been misdiagnosed with schizophrenia and thus have not received an effective therapy.

Sometimes interventions in the brain cause relatively fast changes in personality and behavior that may be reversible. After right pallidotomy for medically treatment-refractory Parkinsonism, a 59-year-old patient developed hypersexuality

including pedophilic behavior. Immediately after the pallidotomy, he became markedly hypersexual. He forced his wife to have sex with him, masturbated frequently, propositioned his wife's female friends, hired strippers and prostitutes, and spent hours viewing Internet pornography. The patient was accused of touching his 5-year-old granddaughter inappropriately and asking her to touch his penis. He was ashamed of his behavior, complained of intrusive sexual thoughts and urges that overwhelmed him, and desired to just have his libidinal urges normalized again. The patient had no history of psychiatric illness, unusual sexual behavior, or drug-induced behavioral changes prior to his surgery. A reduction of his dopaminergic drugs resulted in a gradual decrease in sexual behaviors but worsened the symptoms of Parkinsonism (Mendez and Shapira 2011).

3.2 Fast Processes with Low Plasticity Potential

Particularly dramatic are fast processes with a low plasticity potential. In these cases, the personality and the behavior of an individual change rapidly, that is, within minutes, hours, or days. In addition, the changes are not reversible. Such cases can occur both by brain disease and by interventions in the brain. Again, the fast development of the lesion increases the likelihood that the behavioral changes are perceived as externally caused; however, the low reversibility potential also increases the likelihood of stigmatization.

Strokes can suddenly affect personality and behavior, often irreversibly. In one case, a 70-year-old man developed hemiballism, persistent hypersexuality, memory and executive dysfunction, and poor judgment after a small stroke involving the nucleus subthalamicus (Absher et al. 2000). Another example is the resection of brain tumors, which can change personality or behavior directly and often irreversibly. Although there is no evidence-based knowledge on the incidence, direction and extent of personality changes after brain tumor resection, several studies reveal a relationship between brain tumor surgery and changes in personality and (moral) behavior: Patients who had brain surgery for tumors have higher degrees of emotional and social dysfunction compared to extra-cerebral neurosurgery patients and terminally ill cancer patients (Andrewes et al. 2003, $n = 69$). Particularly, tumor resections from the frontal lobes can cause a lack of emotion and problems with decision-making, even in case of intact cognitive functions. In severe cases, psychopathy can develop which is characterized by impulsivity, antisocial behavior, and uncontrollable aggressions (Phineas Gage syndrome; Damasio 1994; Eslinger and Damasio 1985; Eslinger et al. 2004; Meyers et al. 1992; Tranel et al. 2002, patient SB-2046). Different behavioral disorders have also been reported after surgery for frontolimbic tumors. By way of example, a patient developed kleptomania and compulsive gambling after removal of a craniopharyngioma. Besides this, he became circumstantial and logorrheic, and displayed hypergraphia and a preoccupation with religious and moral ideas (Nyffeler and Regard 2001). In children, aggressive microsurgery for craniopharyngiomas has a significant impact

on social–emotional and behavioral functioning (Sands et al. 2005). A prospective study reports that the majority of children who had total resections of craniopharyngiomas were more or less severely affected by a hypothalamic syndrome that altered their social integration and caused academic failure (Pierre-Kahn et al. 2005, $n = 14$). High rates of intellectual impairment, poor social adaptation, and emotional lability in craniopharyngioma survivors (30–60 %) might be caused by an impaired frontal lobe function following surgery (Stelling et al. 1986).

In some cases, the resection of brain tumors can cause the onset of new psychiatric symptoms. For example, a patient with no previous mental illness developed major depression with psychosis after resection of a giant middle fossa hemangiopericytoma (Sade et al. 2006). Another patient developed a schizophreniform psychosis after excision and postoperative radiotherapy of an oligodendroglioma (Mace and Trimble 1991, case C).

Tumor resections from brain areas, which have recently been considered irrelevant for cognitive capacities, personality, and behavior, can cause a wide spectrum of neuropsychological and behavioral abnormalities. Behavioral deficits or attention deficit problems were detected in 33 % or 12.5 % of patients, respectively, who were operated for benign cerebellar tumors during childhood (Steinlin et al. 2003, $n = 24$). Some demonstrated psychiatric symptoms such as mutism, addiction problems, anorexia, uncontrolled temper tantrums and phobia. Patients had difficulties in selective and, more notably, sustained attention, which resemble dysfunctions seen in patients with frontal lesions (Steinlin et al. 2003). There is evidence that cerebellar dysfunction includes a mild frontal dysfunction due to destroyed cerebello-frontal connections (Steinlin et al. 2003). The resection of benign cerebellar tumors causes the posterior fossa syndrome in 28 % of children, which is characterized by mutism, oropharyngeal dyspraxia, emotional lability, different neuropsychiatric symptoms, and autistic behavior (Catsman-Berrevoets and Aarsen 2010, $n = 148$).

Hypersexuality following brain surgery has also been reported. Two patients developed hypersexuality with inappropriate sexual behavior following the placement of ventriculoperitoneal shunts for the treatment of hydrocephalus (Gorman and Cummings 1992). The sexual disorder was likely caused by septal damage due to the shunt placement. Both lesions and stimulation of the septum have caused hypersexuality in animals and humans. The septum is presumably one locus of a circuit of structures mediating sexual behavior (Gorman and Cummings 1992). Further loci of this circuit include the inferior frontal cortex, the hypothalamus, and the amygdaloid nuclei. Lesions in any of these regions have a major, site-specific impact on sexual behavior. Bilateral lesions of the amygdaloid nuclei produce hypersexuality (Klüver Bucy syndrome); whereas lesions in the hypothalamus reduce sexual activity (Gorman and Cummings 1992). A few long-lasting cases of disinhibition and inappropriate sexual behavior following pallidotomy (lesion of parts of the globus pallidus) for treating medically treatment-resistant Parkinsonism have been published (Shannon et al. 1998, $n = 3$, persistent for at least 6 months).

Fast negative changes in personality and behavior can also occur after psychiatric neurosurgery, which is used very rarely for the treatment of severe therapy-resistant

cases. By way of example, possible sequelae of capsulotomy include aggressiveness, dysexecutive function deterioration, and sexual disinhibition (Cosgrove and Rauch 1995; D'Astous et al. 2013; Dougherty et al. 2002; Feldman et al. 2001; Rück et al. 2008). Following subcaudatetractotomy, the development of undesirable personality traits has been reported in some patients (Feldman et al. 2001). After ventromedial frontal leukotomy, most patients with lesions in the ventral striatum (8 out of 11) developed substance dependence (Irle et al. 1998). The reports indicated no potential of reversibility of these sequelae.

3.3 Slow Processes with High Plasticity Potential

When changes in personality and behavior occur gradually due to slowly developing brain lesions, they allow for better adaptation to the changes, both for the patient and for people in the social surrounding. Slow processes could be considered as less dramatic than fast processes. On the other hand, slow development can conceal the fact that problematic changes in personality and behavior are caused by a disease and not by the patients' "evil will." Particularly, if the disease is not yet diagnosed, or if the patient's significant others do not understand its effects on the patient's personality and behavior, the patient may be blamed for aberrant behavior. However, this risk is diminished if these slowly developing changes in personality and behavior are reversible by adequate medical or neurosurgical therapies. Notable examples can be found particularly in studies about the neurosurgical treatment of epilepsy, Parkinson's disease, and brain tumors.

In patients suffering from epilepsy, changes in personality and behavior mostly develop over several years. However, in many cases, they are reversible after neurosurgical resection of the epileptic focus. Patients with epilepsy have a higher prevalence of lifetime psychiatric disorders (35 %) than the general population (20.7 %; Téllez-Zenteno et al. 2007), and particularly high are the rates in patients with temporal lobe epilepsy (Foong and Flugel 2007). Following surgery for epilepsy, depression, anxiety, behavioral disorders, and severe obsessive–compulsive disorders—which are often, but not always comorbidities of the disease—are often improved (Devinsky et al. 2005; Guangming et al. 2009; Guarnieri et al. 2005; Hannan et al. 2009; Lendt et al. 2000; Witt et al. 2008; review: Foong and Flugel 2007). In many patients, increased warmth in social relationships and reduced egotism have been described (Hill et al. 1957). Improvements in aggressive behavior in children following surgery for temporal lobe epilepsy have been reported in several papers (review: Foong and Flugel 2007). In children, the most notable improvements after surgery for epilepsy include decreased hyperactivity, greater emotional well being, and improved socialization (review: Spencer and Huh 2008).

In patients suffering from Parkinson's disease, some disease-related changes in personality are sometimes reversed after deep brain stimulation (DBS) of the nucleus subthalamicus (Schneider et al. 2003). Changes in behavior caused by

Parkinsonism drug therapy, such as impulse control disorders, pathological gambling, addiction to levodopa, and hypersexuality, can disappear after DBS since it allows for the reduction of the dopaminergic drugs (Demetriades et al. 2011).

Whether patients suffering from brain tumors develop changes in personality and behavior depends on tumor location, tumor size, and tumor type. In many cases, these changes are reversible after resection of the tumor, irradiation, or chemotherapy. The reversibility of tumor-related personality changes depends both on the healing processes in the damaged brain tissue and on the amount of brain tissue that is further damaged by treatment. A notable case in the discussion to follow illustrates how detrimental personality changes can result from a brain tumor and how brain surgery can restore the personality. A 40-year-old married schoolteacher became obsessed with child pornography and started to solicit prostitutes and to molest his stepdaughter. His wife evicted him from the family home after discovering his sexual advances to her daughter. He was accused and found guilty of molesting children. He had to enter a treatment program for convicted sexual offenders where he continued asking women for sex and was expelled from the program. One day before the start of his prison sentence, he was admitted to the hospital for headaches and an indomitable sex drive. An MRI scan revealed that he had an egg-sized brain tumor in the frontal lobe, a brain area essential for judgment, social behavior, and self-control. The tumor had already infiltrated the hypothalamus, which also controls sex drive. After tumor resection, the pedophiliac drive vanished completely, and the patient went home to his family. But several months later, he secretly started to collect pornography again. An MRI scan showed that the tumor had regrown. It was removed once more, and his abnormal sexual drive vanished again (Burns and Swerdlow 2003).

3.4 Slow Processes with Low Plasticity Potential

When changes in personality and behavior occur gradually due to slowly developing brain lesions that in addition have low plasticity potential, the risk that these changes are attributed to the "evil will" of the patient resurfaces, and thus, the problem of misunderstanding, blame, and stigmatization will be aggravated.

Many neurodegenerative brain diseases affect an individual's personality and behavior deeply and enduringly, particularly frontotemporal or vascular dementia, Parkinson's disease and Huntington's disease. Frontotemporal dementia (FTD) is a notable example of sociopathic behavior caused by an acquired frontal brain disorder. FTD is a progressive neurodegenerative disorder that affects mainly the ventromedial prefrontal cortex, anterior temporal regions, and/or the orbitofrontal cortex. The behavioral changes of FTD patients seem to correlate with a decreased metabolism or perfusion in these regions as measured with PET or SPECT (Mendez 2010). FTD patients show marked impairments in moral reasoning despite knowledge of moral and social rules, emotion recognition particularly for anger and disgust, empathy as rated by caregivers, and executive functions (Lough et al. 2006).

FTD patients undergo a change in personality and behavior which is characterized by the following core features: transgression of social norms, sociopathic behavior, altered moral feelings, loss of emotional empathy, and disinhibited, compulsive acts. Although cognition remains largely intact, knowledge of moral behavior and of potential consequences of rule violations is preserved (Mendez 2010). Although they can make reasoned moral judgments, the emotional morality of FTD patients is altered, so that they respond to moral dilemmas in a calculated rather than an emotional fashion (Mendez and Shapira 2009). Typical examples of behavioral problems of FTD patients are loss of social tact and propriety, unacceptable physical contact, neglect of personal hygiene, and compulsive eating or hoarding. More than half of patients fail to conform to lawful behavior. Several reports describe stealing, unethical job conduct, indecent exposure, and inappropriate sexual behavior such as child molestation, illegal driving acts, and physical assaults or violence (Mendez 2010). According to Mendez, FTD patients have a "specific, brain-based impairment in moral reasoning"; their "sociopathic behavior is consistent with decreased emotional moral judgment plus a lack of empathy and disinhibited, compulsive drives" (Mendez 2010, p. 322).

In addition, the growth of brain tumors can affect personality and behavior on a longer timescale. Both tumor type and tumor location are significant influential factors for emotional and social dysfunctions such as anger, helplessness, fatigue, emotional dyscontrol, indifference, and maladaptive behavior (Andrewes et al. 2003, $n = 69$). Tumors in the temporal lobes can be associated with behavioral problems, including aggression and rage attacks (Nakaji et al. 2003). The behavior of patients with lesions in the orbitofrontal and ventromedial PFC has been described as aggressive, lacking responsibility, and concern for social and moral rules (Eslinger and Damasio 1985; Damasio 1994).

4 Ethical Pitfalls in Investigating Changes in Moral Behavior After Brain Lesions

So far, we have outlined the methodological problems when relating moral agency with the brain and provided a scheme to describe the complexity of brain lesion types and moral behavior changes that may result from these lesions. We now discuss the extent to which the investigation of changes in moral behavior after brain lesions poses ethical problems.

Why should it be ethically problematic to relate brain lesions with, in most cases, unwanted changes in moral behavior? This may sound like an odd question, because such findings might contribute to novel therapies that prevent or reverse behavior changes. However, the impetus to find a neuronal cause for disturbed moral behavior aligns with a general tendency in popular culture to find brain-based explanations for behavior (Frazzetto and Anker 2009). Researchers and clinicians who investigate how changes in the brain lead to changes in moral behavior

propose to use their findings for interventions. For example, Fumagalli and Priori (2012, p. 2017) write: "From a clinical point of view, subjects manifesting abnormal moral behavior should be screened for neurological disorders to promote an early diagnosis. A potentially important issue arises when clinicians treat patients whose social position makes them responsible for others (including state leaders and politicians) with abnormalities of moral behavior or with other conditions (or treatment) that could influence their "moral brain." In these cases, an early diagnosis and, whenever possible, effective treatment is important both for the patient and for the welfare of society."

We think that the following questions should be considered in order to assess the ethics of moral behavior interventions based on findings in neuroscience:

1. On the grounds of which ethical theory should the borders between still tolerable behavior and morally blameful behavior be defined?
2. Does it make a difference when a change in moral behavior is caused by a brain lesion due to a pathological process, compared to a change in moral behavior that results as unwanted but maybe inevitable side effect of brain interventions to treat neurological disorders?
3. What do we owe persons who display unpleasant or even immoral behavior due to brain lesions?
4. Does tightening the link between brain damage and behavior aberration increase or decrease the stigmatization of these persons?
5. Given that there is a relation between the brain and moral behavior, what knowledge do we need to better understand this relation?
6. Is there a danger that we pathologize unwanted but legitimate moral standpoints?

In the following, we will briefly discuss these questions and outline some ethical risks that are associated with them.

4.1 Which Ethical Theory?

The first question relates to the basic problem that there is disagreement on what counts as moral behavior. Although we do not support moral relativism, we agree with its observation that moral issues are evaluated very controversially and that the controversies depend significantly on cultural and societal differences. A notable example is the field of sexual morality, where the controversy is enormous. Practices such as prostitution, child marriage, intermarriage, homosexuality, premarital sex, extramarital sex, promiscuity, divorce, polygamy, etc., are socially fully accepted in some cultures and condemned or even illegal in others. This variability may explain the large prevalence differences of hypersexuality, from 2 to 10 % (Chiang et al. 2012) when comparing different countries. The cultural differences in sexual morality probably influence which criteria are used for the diagnosis of hypersexuality and thus the prevalence rates.

For a scientific investigation of brain disorders that cause disorders of moral behavior, a universalistic ethical approach would be optimal, as many ethical theories consider universalizability to be a distinguishing feature of moral judgments and a substantive guide to moral obligation: Moral imperatives should be regarded as equally binding on everyone. However, in philosophy, many arguments have been developed against the feasibility of a universalistic ethic. For example, Beauchamp and Childress' principle-based ethics (2013) that is often considered to be a gold standard in bioethics is exposed to critics from several sides. First, the deductivists (e.g., Clouser and Gert 1990) criticize eclecticism and the lack of a universal, applicable ethical theory. Second, defendants of casuistic ethics (e.g., Jonsen 1995) criticize a too schematic application of principles to particular cases (Harris 2003), and claim that it blocks substantive ethical inquiry (Callahan 2003). Third, the social science critique of bioethics claims that bioethics grounded in philosophy and moral theory gives a dominant role to idealized, rational thought and tends to exclude social and cultural factors, so that it is isolated from practice (overview: Hedgecoe 2004). The dominance of the principle of respect for autonomy in particular has been criticized by many authors from different ideological backgrounds (critical overview: Gillon 2003). In spite of the diversity of these criticisms, they converge in giving collective benefits more weight than individual rights.

In summary, the question of "which morality?" does not only refer to the commonplace, that there is disagreement concerning the morality of certain behaviors. The point is that the relationship of brain lesions with moral behavior tends to blur this variability and that this may happen on a level where this effect is difficult to be discerned, for example, on the level of diagnostic criteria. This impetus to universalize morality is not based on grounds of an elaborated theory of ethical universalism, but is implicitly embedded in the methodology that investigates the phenomenon. Therefore, we identify as the first ethical risk of the neuroscience of ethics that it may suppress legitimate controversies on moral theories.

4.2 What Is Causing the Brain Lesion?

At first glance, there seems to be a fundamental difference between changes of moral behavior caused by pathological processes such as strokes or tumors, and those caused by medical interventions. This distinction may hold when the intervention directly targets the behavior as in psychiatric neurosurgery, but it is less clear for lesions that occur as unintended or unavoidable consequences of therapeutic interventions such as tumor resection.

For dealing with this problem, it is necessary to accumulate knowledge on sequelae that may result from particular interventions. This knowledge then can be used for the shared decision-making process between patient and medical experts. The problem, however, is that changes in moral behavior that are caused by

interventions are often hard to measure, whereas their relative life impact is high (Müller and Christen 2011).

However, a fundamental problem remains, namely the large individual variability of human brains (e.g., functional connectivity; Barch et al. 2013) as well as of regeneration processes. Therefore, clinical outcome studies that average across patients to provide a unitary measure of outcome are not sufficient. Because of the large outcome variability, it is necessary to report both good and poor outcomes separately. Cross-sectional group research does not reveal the different individual trajectories and provides only limited clues about which factors are most relevant in effecting positive change for an individual. It is important to study individual outcomes, particularly by identifying subgroup patterns that can become lost in whole-group analyses. To overcome this systematical shortcoming, long-term follow-up studies of outcome, particularly of neuropsychological and socio-psychological outcome, are necessary (Wilson et al. 2005). Particularly cases with unfavorable or unexpected outcome should be investigated, since they offer extraordinary chances for scientific discovery and improving the techniques used (Kubu and Ford 2012). Besides clinical studies, case studies contribute much to clinical experience and to scientific understanding. For example, the knowledge of adverse effects of deep brain stimulation has been spread mainly via case reports (Christen and Müller 2011). The careful documentation and publication of extraordinary single cases are important for scientific progress. This highlights the importance of case studies in addition to knowledge based on statistical evidence.

Given this caveat, an approach in therapy and rehabilitation that focuses on the individual case seems appropriate. However, this approach conflicts with an increasing involvement of knowledge based on statistical evidence and a regulatory or legal structure that more and more relies on such knowledge, for example, in addressing insurance claims. This pinpoints a second ethical risk: Restricting clinical research on the relationship between brain and moral behavior may undermine the value of special case studies involving outliers.

4.3 What Do We Owe Brain Lesion Patients with Socially Aberrant Behavior?

The care of frontal lesion patients is challenging and demonstrates constraints of classical principles of medical ethics like autonomy and beneficence. Disabilities that directly affect social interactions with others pose more challenges to family and caregivers compared to physical disabilities or pure cognitive disabilities. The main reason for this is probably that the disability caused by the brain lesion directly influences morally questionable behaviors like boasting, egocentrism, or obstinacy; or even uncontroversial immoral behaviors like habitual lying, child molestation, or violence. Empathy or even sympathy for these patients is much harder to sustain.

Most people who suffer from brain lesions are not responsible for their lesions, and they are significantly disadvantaged. Therefore, we think that society has the moral duty to support their reintegration. Besides a good medical treatment, rehabilitation programs are necessary, and for some patients a protected environment. Furthermore, we think that research on the responsibility of people with brain lesions for aberrant or even criminal behavior is necessary. This research has to consider both medical and normative issues and therefore requires interdisciplinary programs.

4.4 Will Research Increase the Stigmatization of Patients with Brain Lesions?

It is tempting to believe that a better understanding of how damage to the brain leads to changes in moral behavior will increase the understanding of such patients and enhance their social reintegration. However, we suspect that this hope is misleading for two reasons. First, experimental ethics has shown that the majority of people judge in an incompatibilist way, that is, they believe that determinism excludes moral responsibility. But many people tend toward compatibilism if the cases they have to judge trigger emotions; then most people blame others and hold them morally responsible despite knowing that the person was determined to act in a specific, immoral way (Nichols and Knobe 2007). Second, research on stigmatization has shown that biological explanations of psychiatric disorders have complex effects on stigmatization which depend on several aspects of a given disorder. Biological explanations of psychiatric disorders increase stigmatization particularly if the disorder is explained as irreversible or as genetically based, or if it makes the patients dangerous for third persons (Müller and Heinz 2013). Therefore, we expect that a better biological understanding of how brain lesions cause aberrant behavior and changes in personality might decrease the stigmatization of principally curable or reversible diseases but increase the stigmatization of irreversible brain disorders.

From a theoretical point of view, a deeper understanding of how brain lesions cause social aberrant behavior is likely to support the view that brain lesion patients also deserve help and understanding when they behave immorally. However, it is likely that for many patients the practical effect will be the opposite—namely more stigmatization and discrimination. Thus, the third ethical risk of a neuroscience of ethics is that a better understanding of how brain lesions cause moral behavior changes might undermine people's willingness to support lesion patients to reintegrate in the society.

4.5 What Should We Investigate?

This leads to the problem of deciding on which side of the interrelation between the brain and moral behavior the emphasis of research should be. Here, we are confronted with a fundamental impetus of the scientific method—namely that it aims to generate causal knowledge that is as precise and as deterministic as possible. The neuroscience of ethics proceeds from documenting correlations between brain lesions and social behavior to experimenting by noninvasive interventions with transcranial magnetic stimulation or by drugs. This should allow for finding possible causal relationships underlying the correlations—for example, that inhibition of a certain brain region causes lack in impulse control. However, as our overview in Sect. 3 has shown, there is a surprising lack of knowledge on the reversibility of such behaviors. We propose that future research in the neuroscience of ethics should focus on three issues: the spectrum of behavior aberrations following brain lesions including their interrelation with the social environment, the chances of regeneration, and the development of effective therapies.

4.6 Danger of Pathologizing Ethical Theories

We also pinpoint a fourth ethical risk of the neuroscience of ethics, namely to relate different ethical standpoints or ethical theories like deontology or utilitarianism to specified neuronal differences, or even to certain neurological disorders. It is possible to detect different activation patterns during different forms of moral thinking. But what is the function of such findings in the meta-ethical discourse? We suspect that they are sometimes misused for debunking ethical theories, that is, for discrediting ethical theories not by normative arguments, but by pinpointing inopportune or emotional processes that cause certain types of moral argumentations. An example is the discretization of deontology compared to utilitarianism based on neuroscientific findings (Singer 2005). In extremis, this could subserve a pathologization of ethical theories.

5 Summary and Outlook: Moral Behavior as Target of Therapy

The central aim of our contribution is to outline the complexity of the interrelation between the brain and moral behavior relation when seen from a neuroscientific standpoint. We have argued that dynamic and temporal factors on both sides of the equation—the stability of moral evaluations across time and society or across neuronal plasticity—structure this complexity, but also explain why we cannot expect clear-cut relations between specific brain lesions and specific behavior

aberrations. But to achieve feasibility for therapy and rehabilitation, we will need models that have an intermediate degree of complexity, like the model of moral intelligence we have proposed. These models must be complex enough to capture the relevant phenomenon, but simple enough to be understandable for practical purposes. In that way, basic researchers and therapists can approach the question of how to interrelate the brain with moral agency in a more structured way.

However, one has to be aware of ethical risks that accompany this endeavor. The most urgent risk is a lack of knowledge about the spectrum of behavioral aberrations that accompany specified brain lesions, or about the potential of regeneration and its dependence on the social environment. The problem is aggravated by the difficulties in everyday life when dealing with patients that show aberrant social behavior, where one has to balance demands for responsible behavior and lenience due to their disability.

We close by remarking that the complexity outlined in our contribution also calls for caution with respect to recent claims for moral enhancement—the idea that knowledge of the biological foundation of human moral behavior may allow for interventions into the neuronal infrastructure of morality in order to improve the behavior of people or, at least, to diminish some forms of evil (De Ridder et al. 2009; Shook 2012; Persson and Savulescu 2012). Although we are optimistic that more sophisticated and individualized research will certainly help to bear the behavioral burdens caused by some brain lesions, we do not support policies of moral enhancement through brain interventions that disregard the autonomy and dignity of the patients concerned.

Acknowledgments Funding: This research has been partly supported by the German Research Foundation (DFG), Germany (MU 3321/1-1).

References

Abend G (2013) What the science of morality doesn't say about morality. Phil Soc Sci 43:157–200

Absher JR, Vogt BA, Clark DG et al (2000) Hypersexuality and hemiballism due to subthalamic infarction. Neuropsych Neuropsychol Behav Neurol 13(3):220–229

Andrewes DG, Kaye A, Murphy M et al (2003) Emotional and social dysfunction in patients following surgical treatment for brain tumour. J Clin Neurosci 10(4):428–433

Appiah KA (2010) The honor code: how moral revolutions happen. W.W. Norton & Company, New York

Barch DM, Burgess GC, Harms MP et al (2013) Function in the human connectome: task-fMRI and individual differences in behavior. NeuroImage 80:169–189

Batts S (2009) Brain lesions and their implications in criminal responsibility. Behav Sci Law 27 (2):261–272

Beauchamp TL, Childress JF (2013) Principles of biomedical ethics, 7th edn. Oxford University Press, Oxford

Björklund F, Björnsson G, Eriksson J et al (2012) Recent work: motivational internalism. Analysis 72:124–137

Burns JM, Swerdlow RH (2003) Right orbitofrontal tumor with pedophilia symptom and constructional apraxia sign. Arch Neurol 60:437–440

Callahan D (2003) Principlism and communitarianism. J Med Ethics 29:287–291
Catsman-Berrevoets C, Aarsen FK (2010) The spectrum of neurobehavioral deficits in the posterior fossa syndrome in children after cerebellar tumour surgery. Cortex 46:933–946
Chiang HL, Huang YS, Chen ST et al (2012) Are there ethnic differences in impulsive/compulsive behaviors in Parkinson's disease? Europ J Neurol 19(3):494–500
Cholbi MJ (2006) Belief attribution and the falsification of motive internalism. Philos Psychol 19:607–616
Christen M, Alfano M (2013) Outlining the field—a research program for empirically informed ethics. In: Christen M, Fischer J, Huppenbauer M, Tanner C, van Schaik C (eds) Empirically informed ethics. Morality between facts and norms. Library of ethics and applied philosophy. Springer, Berlin, pp 3–28
Christen M, Müller S (2011) Single cases promote knowledge transfer in the field of DBS. Front Integr Neurosci 5 (Article 13)
Christen M, Regard M (2012) Der "unmoralische Patient". Eine Analyse der Nutzung hirnverletzter Menschen in der Moralforschung. Nervenheilkunde 31:209–214
Ciaramelli E, Muccioli M, Làdavas E et al (2007) Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. Soc Cog Affect Neurosci 2:84–92
Clouser KD, Gert B (1990) A critique of principlism. J Med Philos 15:219–236
Cosgrove GR, Rauch SL (1995) Psychosurgery. Neurosurg Clin North A 6:167–176
D'Astous M, Cottin S, Roy M et al (2013) Bilateral stereotactic anterior capsulotomy for obsessive-compulsive disorder: long-term follow-up. J Neurol Neurosurg Psychiat 84 (11):1208–1213
Dalmau J, Gleichman AJ, Hughes EG et al (2008) Anti-NMDA-receptor encephalitis: case series and analysis of the effects of antibodies. Lancet Neurol 7:1091–1098
Dalmau J, Tuzun E, Wu HY et al (2007) Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with ovarian teratoma. Ann Neurol 61:25–36
Damasio AR (1994) Descartes' error. Putnam, New York
Demetriades P, Rickards H, Cavanna AE (2011) Impulse control disorders following deep brain stimulation of the subthalamic nucleus in Parkinson's disease: clinical aspects. Parkinson's Disease, doi: 10.4061/2011/658415
De Ridder D, Langguth B, Plazier M et al (2009) Moral Dysfunction and potential treatments. In: Verplaetse J et al (eds) The moral brain. Springer, Berlin, pp 155–183
Devinsky O, Barr WB, Vickrey BG et al (2005) Changes in depression and anxiety after resective surgery for epilepsy. Neurology 65:1744–1749
Dougherty DD, Baer L, Cosgrove GR et al (2002) Prospective long-term follow-up of 44 patients who received cingulotomy for treatment-refractory obsessive-compulsive disorder. Am J Psych 159:269–275
Eslinger PJ, Damasio A (1985) Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. Neurology 35:1731–1741
Eslinger PJ, Flaherty-Craig CV, Benton AL (2004) Developmental outcomes after early prefrontal cortex damage. Brain Cog 55:84–103
Feinstein JS, Rudrauf D, Khalsa SS et al (2010) Bilateral limbic system destruction in man. J Clin Exp Neuropsychol 32(1):88–106
Feldman RP, Alterman RL, Goodrich JT (2001) Contemporary psychosurgery and a look to the future. J Neurosurg 95:944–956
Foong J, Flugel D (2007) Psychiatric outcome of surgery for temporal lobe epilepsy and presurgical considerations. Epilepsy Res 75:84–96
Frazzetto G, Anker S (2009) Neuroculture. Nature Rev Neurosci 10(11):815–821
Frías Ibáñez A, Malea Fernández A, Dorado García M et al (2008) Orbitofrontal syndrome after a cerebral aneurysm operation. A case report. Actas España Psiquiatrica 36(6):362–365
Fumagalli M, Priori A (2012) Functional and clinical neuroanatomy of morality. Brain 135 (7):2006–2021
Gillon R (2003) Ethics needs principles—four can encompass the rest—and respect for autonomy should be "first among equals". J Med Ethics 29:307–312

Gorman DG, Cummings JL (1992) Hypersexuality following septal injury. Arch Neurol 49:308–310
Grafman J, Schwab K, Warden D et al (1996) Frontal lobe injuries, violence, and aggression: a report of the Vietnam Head Injury Study. Neurology 46:1231–1238
Guangming Z, Wenjing Z, Guoquiang C et al (2009) Psychiatric symptom changes after corticoamygdalo-hippo-campectomy in patients with medial temporal lobe epilepsy through Symptom Checklist 90 Revised. Surgical Neurol 72:587–591
Guarnieri R, Araújo D, Carlotti CG Jr et al (2005) Suppression of obsessive-compulsive symptoms after epilepsy surgery. Epilepsy Behav 7:316–319
Hannan S, Cross JH, Scott RC et al (2009) The effects of epilepsy surgery on emotions, behavior, and psychosocial impairment in children and adolescents with drug-resistant epilepsy: a prospective study. Epilepsy Behav 15:318–324
Harlow JM (1868) Recovery from the passage of an iron bar through the head. Publ Mass Med Soc 2:327–347
Harris J (2003) In praise of unprincipled ethics. J Med Ethics 29:303–306
Hawkins KA, Trobst KK (2000) Frontal lobe dysfunction and aggression: conceptual issues and research findings. Aggr Viol Behav 5(2):147–157
Hedgecoe AM (2004) Critical bioethics: beyond the social science critique of applied ethics. Bioethics 18(2):120–143
Hill D, Pond DA, Mitschell W et al (1957) Personality changes following temporal lobectomy for epilepsy. J Med Sci 103:18–27
Hornak J, Bramham J, Rolls ET et al (2003) Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. Brain 126:1691–1712
Irle E, Exner C, Thielen K et al (1998) Obsessive-compulsive disorder and ventromedial frontal lesions: clinical and neuropsychological findings. Am J Psych 155:255–263
Jonsen AR (1995) Casuistry: an alternative or complement to principles? Kennedy Inst Ethics J 5:237–251
Koenigs M, Young L, Adolphs R et al (2007) Damage to the prefrontal cortex increases utilitarian moral judgments. Nature 446:908–911
Kohlberg L (1981/1984) Essays on moral development, vol I/II. Harper & Row, San Francisco
Kubu CS, Ford PJ (2012) Beyond mere symptom relief in deep brain stimulation: an ethical obligation for multi-faceted assessment of outcome. AJOB Neurosci 3(1):44–49
Lendt M, Helmstaedter C, Kuczaty S et al (2000) Behavioral disorders in children with epilepsy: early improvement after surgery. J Neurol Neuro-surg Psych 69:739–744
Lough S, Kipps CM, Treise C et al (2006) Social reasoning, emotion and empathy in frontotemporal dementia. Neuropsychologia 44(6):950–958
Mace CJ, Trimble MR (1991) Psychosis following temporal lobe surgery: a report of six cases. J Neurol Neurosurg Psych 54:639–644
Macmillan M (2000) An odd kind of fame. Stories of Phineas Gage. MIT-Press, Cambridge, Mass
Mataró M, Jurado MA, García-Sánchez C et al (2001) Long-term effects of bilateral frontal brain lesion: 60 years after injury with an iron bar. Arch Neurol 58(7):1139–1142
Mendez MF, Shapira JS (2011) Pedophilic behavior from brain diseases. J Sex Med 8:1092–1100
Mendez MF (2010) The unique predisposition to criminal violations in frontotemporal dementia. J Am Acad Psych Law 38(3):318–323
Mendez MF (2009) The neurobiology of moral behavior: review and neuropsychiatric implications. CNS Spectrum 14(11):608–620
Mendez MF, Shapira JS (2009) Altered emotional morality in frontotemporal dementia. Cogn Neuropsych 14(3):165–179
Meyers CA, Berman SA, Scheibel RS et al (1992) Acquired antisocial personality disorder associated with unilateral left orbital frontal lobe damage. J Psych Neurosci 17(3):121–125
Müller S, Heinz A (2013) Stigmatisierung oder Entstigmatisierung durch Biologisierung psychischer Krankheiten? Nervenheilkunde 12:955–961
Müller S, Christen M (2011) Deep brain stimulation in Parkinsonian patients—ethical evaluation of cognitive, affective, and behavioral sequelae. AJOB Neurosci 2(1):3–13

Nakaji P, Meltzer HS, Singel SA et al (2003) Improvement of aggressive and anti-social behaviour after resection of temporal lobe tumors. Pediatrics 112(5):e430–e433
Narvaez D (2005) The neo-Kohlbergian tradition and beyond: schemas, expertise and character. In: Carlo G, Pope-Edwards C (eds) Moral motivation through the lifespan. University of Nebraska Press, Lincoln, pp 119–163
Nichols S, Knobe J (2007) Moral responsibility and determinism: the cognitive science of folk intuitions. Noûs 41(4):663–685
Nyffeler T, Regard M (2001) Kleptomania in a patient with a right frontolimbic lesion. Neuropsych Neuropsychol Behav Neurol 14(1):73–76
Parkinson C, Sinnott-Armstrong W, Koralus PE et al (2011) Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. J Cog Neurosci 23(10):3162–3180
Pascual L, Rodrigues P, Gallardo-Pujol D (2013) How does morality work in the brain? A functional and structural perspective of moral behavior. Front Integr Neurosci 7:65
Persson I, Savulescu J (2012) Unfit for the future: the need for moral enhancement. Uehiro series in practical ethics. Oxford University Press, Oxford
Piaget J (1932) The moral judgment of the child. Kegan Paul, Trench, Trubner and Co., London
Pierre-Kahn A, Recassens C, Pinto G et al (2005) Social and psycho-intellectual outcome following radical removal of craniopharyngiomas in childhood: a prospective series. Child´s Nerv Sys 21:817–824
Rest JR (1986) Moral development: advances in research and theory. Praeger, New York
Roskies A (2006) Patients with ventromedial frontal damage have moral beliefs. Philos Psychol 19:617–627
Rück C, Karlsson A, Steele D et al (2008) Capsulotomy for obsessive-compulsive disorder. Arch Gen Psych 65(8):914–922
Sade B, Budur K, Lee DK et al (2006) Major depression with psychosis after resection of a giant middle fossa hemangiopericytoma. Surg Neurol 65:290–292
Sands SA, Milner JS, Goldberg J et al (2005) Quality of life and behavioral follow-up study of pediatric survivors of craniopharyngioma. J Neurosurg 103(4 suppl):302–311
Schneider F, Habel U, Volkmann J et al (2003) Deep brain stimulation of the subthalamic nucleus enhances emotional processing in Parkinson disease. Arch Gen Psych 60(3):296–302
Shamay-Tsoory SG, Tomer R, Goldsher D et al (2004) Impairment in cognitive and affective empathy in patients with brain lesions. J Clin Exp Neuro-psycho-l 26(8):1113–1127
Shannon KM, Penn RD, Kroin JS et al (1998) Stereotactic pallidotomy for the treatment of Parkinson's disease. Efficacy and adverse effects at 6 months in 26 patients. Neurology 50:434–438
Shook JR (2012) Neuroethics and the possible types of moral enhancement. AJOB Neurosci 3 (4):3–14
Singer P (2005) Ethics and intuitions. J Ethics 9:331–352
Spencer S, Huh L (2008) Outcomes of epilepsy surgery in adults and children. Lancet Neurol 7:525–537
Steinlin M, Imfeld S, Zulauf P et al (2003) Neuropsychological long-term sequelae after posterior fossa tumour resection during childhood. Brain 126:1998–2008
Stelling MW, McKay SE, Carr A et al (1986) Frontal lesions and cognitive function in craniopharyngioma survivors. Am J Dis Child 140:710–714
Tanner C, Christen M (2013) Moral intelligence—a framework for understanding moral competences. In: Christen M, Fischer J, Huppenbauer M, Tanner C, van Schaik C (eds) Empirically informed ethics. Morality between facts and norms. Library of ethics and applied philosophy. Springer, Berlin, pp 119–136
Téllez-Zenteno JF, Dhar R, Hernandez-Ronquillo L et al (2007) Long-term outcomes in epilepsy surgery: antiepileptic drugs, mortality, cognitive and psychosocial aspects. Brain 130:334–345
Tranel D, Bechara A, Denburg NL (2002) Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. Cortex 38:589–612

Uttal WR (2001) The new phrenology. The limits of localizing cognitive processes in the brain. MIT Press, Cambridge

Varadkar S, Bien CG, Kruse CA et al (2014) Rasmussen's encephalitis: clinical features, pathobiology, and treatment advances. Lancet Neurol 13(2):195–205

Welt L (1888) Über Charakterveränderungen des Menschen infolge von Läsionen des Stirnhirns. Deutsches Archiv für klinische Medicin 42:339–390

Wilson SJ, Bladin PF, Saling MM et al (2005) Characterizing psychosocial outcome trajectories following seizure surgery. Epilepsy Behav 6:570–580

Witt JA, Hollmann K, Helmstaedter (2008) The impact of lesions and epilepsy on personality and mood in patients with symptomatic epilepsy: a pre- to postoperative follow-up study. Epilepsy Res 82:139–146