Chapter 1

Invoking the brain in studying morality: A theoretical and historical perspective on the neuroscience of morality

Published as:

Abstract
Contemporary neuroscientific approaches try to make morality accessible and assessable by linking it to the brain. This wedding between a material entity and an elusive, multifaceted notion has both distinctive preconditions and complex implications. The triadic framework of locating, translating, and defining morality in terms of the brain is applied to illustrate the process of utilizing the cerebral realm for studying the moral realm. A historical and theoretical perspective is introduced to investigate how the relation between brain and morality has been and is being established. It is demonstrated that the meaning of both the brain and morality is in constant flux and dependent upon theorizing and research practice. Thus, I argue that what the neuroscience of morality studies is a contingent notion.
“Morality is as firmly grounded in neurobiology as anything else we do or are. Once thought of as purely spiritual matters, honesty, guilt, and the weighing of ethical dilemmas are traceable to specific areas of the brain.” (de Waal, 1996 pp. 217 - 218)

The brain is pervasive. It affords every aspect of our psychological life, including morality. Given this essential function, it prevails in the study of humans. It coaxes researchers towards studying it as a source of mind and behavior, covering topics from romantic love to cultural conflict (Gergen, 2010). It catches researchers’ attention by luring them into thinking that the human subject is in fact a cerebral subject. It endorses the popular doctrine that it will reveal itself through scrupulous examination and thereby furnish us humans with seminal knowledge about ourselves (Vidal, 2009). The brain is powerful indeed; and as the soaring amount of papers on the neuroscience of morality testifies, it is being invoked in studying morality. Advocates of this undertaking postulate that the origins, causes, and concomitants of moral mind and behavior are best to be understood by investigating the composition and functioning of the brain (for reviews see Fumagalli & Priori, 2012; Mendez, 2009).

What is this powerful thing in our head that allegedly drives our moral conduct? Is it the fixed material entity that neuroscience takes it to be? The history of brain research amply illustrates how the brain, as an object of study, has been subject to change. Whereas the Ancient Egyptians ignored it, modern neuroscience idolizes it (Finger, 2005; Gazzaniga, 2005; Hagner, 2008). Thinking about the brain and ascribing attributes to it is of long standing. Assumptions about its constitution and capacity have varied over time. What are the implications of the brain’s changing history? Could it be that the brain is not rigid, but in flux, and, as a consequence, evasive? To specify: What is changing is not the brain per se – that is, the material entity as it is putatively studied by neuroscience – but the notion of the brain – that is, the specific
set of assumptions, methods, theories, and practices that researchers utilize and adhere to in order to investigate it. What is thought of the brain and what it “really” is are impossible to keep apart clearly. Correspondingly, the notion of the brain is and was being established in different contexts of use within diverse communities, resulting in a multifarious object of study that is flexible and contingent to a time and place (Abi-Rached & Rose, 2010).

Proclaiming the vicissitude of the notion of the brain relativizes its role as a pervasive mediator by putting it into historical perspective. Furthermore, suggesting that neuroscientists who conceive of themselves as studying a material entity in fact study a volatile set of notions obviously is a strong claim against the realism interweaved in their endeavor. The point here is not to dispute the brain’s possible significance in the generation of human mind and behavior, or to vilify neuroscience, but to show that what modern neuroscience studies is a specific notion of the brain which has been (and most likely will be) subject to change in the course of history. Accordingly, researching and understanding morality via the brain are subject to change, dependent upon the given status of neuroscience. That the meaning of morality is ambiguous – that is, its equivocality invites versatile definitions (Gert, 2011) – contributes to its malleability in terms of the brain. Correspondingly, I use the word “morality” here and throughout this essay in a broad and ambiguous sense. Keeping this in mind, the goal is to demonstrate that the given corpus of scientists’ assumptions about the brain has influenced the conceptualization of morality as an object of the neurobiological sciences. The result is an idiosyncratic determination of morality in terms of the brain. In order to show this, I will analyze the process of utilizing the brain for the study of morality via the triadic framework of locating, translating, and defining. Subsequently, the history of the combined study of brain and morality in the 20th century will be sketched to show the inconstancy of their
relation over time. In conclusion, the specific concept of morality that is studied by modern neuroscience is explored.

**Making the brain the moral domain**

According to some scholars who are enthusiastic about neuroscience, morality is to be transformed into “a problem of brain biology” (Tancredi, 2005 p. x). This implies that the fabric of our moral conduct is not individual experiences, not social relations, not cultural traditions, but neurons and axons. The brain is the moral domain. Before the establishment of this domain is analyzed, some clarification on the tacit notion of the brain in neuroscience is required here.

First, neuroscientific approaches to human mind and behavior in general and to morality in particular rest on the assumption that the brain is a reliable informant. Its reliability is assumed owing to its putative impartiality: the brain does not have an opinion (unlike biased ethics professors); it simply provides facts. The brain as part of the natural realm is perceived to be a neutral arbitrator: it is non-human, thus non-subjective, and hence by studying it objective knowledge about morality can be derived. After all, how could nature be wrong? Second, the assumption of neutrality extends to the idea of historical stability: it is and always has been the same thing that is studied. Third, and perhaps most importantly, the brain appeals because it promises a solid foundation of morality in the natural realm. Individuals, groups, and cultures are fickle; the brain is not. Its firm biological roots in nature add credibility to claims derived from it.

But how is this bundle of neurons employed to convert morality into “a problem of brain biology”? The conversion is a complex process of utilization that entails a plethora of philosophical, methodological, and empirical questions (Kahane & Shackel, 2010; N. Levy, 2007; Schleim & Schirmann, 2011). Here, the focus is on
forming the relation between the brain and morality. Consider the following abbreviated quotation from Tancredi (2005 pp. 9-10):

“Three of the Seven Deadlies [si] […] have already been shown to be affected by biological factors in varying degrees, and in some cases the individual may have little power, or “free will,” to prevent them from happening. […] With regard to the most conspicuous display of society’s notion of gluttony – obesity – […] [m]any causes of obesity have nothing to do with lack of control or excessive self-indulgence. Metabolism, which is genetically determined, can result in weight gain despite efforts toward control of excessive eating. Studies have shown that obesity may involve either of two brain systems: the system that sends hunger and satiation messages to the brain or the system associated with the reward circuits involved in drug […] addiction. […] Research is already showing that gluttony and obesity may involve abnormalities in specific areas of the brain. […] [T]he human brain is highly sensitive to food and [research has shown] that the presence of food increases brain metabolism in specific areas. Increased metabolism in the right orbitofrontal cortex correlates highly with self-reports of increased hunger and desire for food.”

When reading this quotation superficially, one is left with the impression that attaching moral notions to the brain is self-evident, long overdue, and scientifically substantiated. The metamorphosis from gluttony to obesity to metabolism to genetic determination to brain involvement goes swiftly and smoothly. However, a closer look reveals that the conversion, and subsequent dissolving, of a Deadly Sin into the functioning of brain areas incorporates three processes: locating, translating, and defining.
Locating refers to the process of transferring gluttony from the verbal and social realm into the physical realm: An indication for gluttony is placed in the “right orbitofrontal cortex.” An amorphous word – that is, a descriptive term for certain behaviors with normative connotations that is embedded in culture and religion – is assigned to a cerebral region, a part of the physical world. A compulsory change of residence is made for the Deadly Sin. The second process, translating, incorporates a re-description of gluttony in terms of bio-scientific language, for example metabolism. This amounts to more than harmless metonymies. Gluttony is classified as a word in the neuroscientific language game. This has the effect that more neuro-language can now be attached and applied to it (“reward circuits involved in drug … addiction”). Gluttony now makes sense in neuroscientific lingo (see Gergen, 2010). The third process is defining. Gluttony is curtailed in order to make it tangible for neuroscience. First, certain attributes of gluttony that bear resemblances to known features of brain functioning are highlighted and isolated. For example, the presence of and sensitivity to food in an experimental design is taken to be indicative of it (see Wang et al., 2004, for the original study). Thus, gluttony is divided into subcomponents that are neuroscientifically measurable and function as indicators for the Deadly Sin. Silently, the rest of gluttony’s connotation, such as indulging in excessive luxury, is dismissed. This rendering is in essence a process of operationalization (see Crawford, 2010; Uttal, 2001, on this issue). In general, an operationalization serves the purpose to specify operations that make the concept at hand measurable. The operationalization here is a particular and intricate one: Gluttony has to be made measurable in terms of the brain, therefore certain subcomponents are emphasized which are (putatively) accessible with neuroscientific methods and compatible with neuroscientific knowledge. As a consequence, gluttony is cropped, split, and diminished in order to be traceable in the cerebral purview (see S. Cohn, 2011). Defining gluttony in such a way generates an epistemological gateway for neuroscience.
After gluttony is fixed in the brain (locating), addressed in terms of the brain (translating), and trimmed in order to fit the constraints of the neuroscientific experiment (defining), its metamorphosis into “a problem of brain biology” is complete. Philosophically speaking, this conversion is a shift in ontology (locating), semantics (translating), and epistemology (defining). This transformation is a precondition for neuroscientists’ quest to trace morality in the brain. Moreover, the process described above needs to be conducted for every aspect of morality individually, since each has its own idiosyncrasies. Other instructive examples are the transformations of aspects of morality, such as moral perception, moral conflict, moral argument, and moral pathology, into neural network functioning (Churchland, 1998, for these examples).

Establishing the brain–morality relation

The example in the previous section demonstrates that some forging needs to be done in order for the brain to testify to morality. The essential step is establishing a relation between the brain and morality in order to constitute the brain as moral domain. The general idea is to connect morality, with all its versatile semantic facets, to an anatomical entity, the brain.

Relating an amorphous notion to a graspable material object appears to be a complicated wedding. In order to avoid complication, three important things need to be kept in mind. First, what are related are notions, which are specific sets of assumptions, methods, theories, and practices. While this might be rather uncontroversial for morality, this also holds true for the brain (see below). Second, the relation between the brain and morality does not propose itself by virtue of nature; it is established via neuroscientific reasoning in a specific way. Third, the relation is not a product, but a process. It is constantly being assessed and manufactured because it is
contingent to what is known of the brain, morality, and the relation of the two. The establishment of this relation is an actual problem that neuroscientists face and need to solve before they can start experimenting. They need to link morality to the brain in order to make the former measurable in terms of the latter. This is essentially a process of granting access (as the gluttony example demonstrated). Neuroscientists’ goal is to make morality empirically accessible and assessable by introducing a material purview, brain matter, which can be studied in an experimental fashion. The intention is to identify patterns – that is, recurring, consistent data – which add credibility to the claim that the brain and morality are associated. Connected with this is a general problem that demands neuroscientists’ attention: Owing to the fact that the above-mentioned connection combines the physical and biological with the psychological, it needs to be based on a general bridge principle that connects the brain with the mind – that is, a substantiated theory that unites the two realms. Numerous such theories have been proposed (e.g., eliminative materialism Churchland, 1981), yet they are subject to fierce debates. Thus, the general brain–mind interplay is contested and, maybe as a consequence, rarely elucidated in the neuroscience of morality literature. In a similar vein, explicit statements regarding the structure of the brain-morality relation are rare. How precisely the brain and morality are supposed to interact remains obscure: Is their relation causal, correlative, epiphenomenal, …?

Additionally, a relation between brain and morality can allow for influences in both directions: the brain could impact morality and/or morality could affect the brain. Historically and contemporarily, the focus of study has been one-sided. Neuroscientists have devoted more attention to the way the brain acts upon morality than vice versa (Sinnott-Armstrong, 2008; Verplaatse, 2009). Thus, with respect to researchers’ interests, it is the brain that comes first and morality (somehow) follows. This is a statement about succession that does not necessarily imply causation. Yet, science has centered on the role of the brain for morality which ranges from
association over mediation, and foundation, to determination. Keeping these clarifications in mind, the following section demonstrates that the brain-morality relation is mutable.

The brain–morality relation in flux

The tethering of brain and morality is a process of biologization which alters the understanding and the study of morality. In the course of history, scientists have assigned various roles to the brain (Vidal, 2009) which had distinct influences on its expressiveness for morality. The understanding of morality as advocated in the different scientific approaches has varied accordingly. As will be explained below, it is noteworthy that the specific type of brain–morality relation of a given time is closely connected to the reigning neuro-/biological theories of the day.

The inception
The science-driven biologization of morality has branched roots. In the 17th and 18th centuries, scholars discussed the meaning of animal instinct for human reasoning and social functioning (Richards, 1987). Others employed the authority of nature and biology to legitimate societal categories (Daston, 1992). Benjamin Rush, one of the founding fathers of American psychiatry, identified physical factors that impacted the moral faculty at the end of the 18th century (Rush, 1786). In the 19th century, Franz Joseph Gall, the champion of phrenology, identified a “moral” organ in the cortex of the brain (Verplaetse, 2009). Around the same time, biomedically minded experts proclaimed immorality to be a somatic disease (Harris, 1989; Rimke & Hunt, 2002) and theories on criminal bodies and brains proliferated (Rafter, 2008).

Moreover, Darwin’s theory of evolution grounded morality in the natural world. According to Darwin, this was a gruesome place in which organisms competed ruthlessly and tried to prevail in overpowering each other. Morality also had to make
sense in this framework. Darwinian theory deemed morality to be inheritable and expressed in instinctive behavior; thus, morality was endowed with biological qualities (Darwin, 1880). Most importantly, morality served a function: it was an advantageous trait in the struggle for survival. It increased the individual organism’s chances to survive and, ultimately, reproduce. As a consequence, genetic makeup for morality was more likely to be passed on to the following generation and by this process proliferated. Consequently, the origin of morality was transferred from a religious source to the history of human evolution. This descriptive statement about the genesis of morality became more and more accepted. However, the possible normative implications were subject to controversy. Comprehensive complaints were voiced against the idea that evolutionary principles should guide conduct in society (e.g. Huxley, 1893).

Darwinian ideas influenced the scientific reasoning about morality at the end of the 19th century (Richards, 1987). The underlying problem that evolution needed a vehicle to transfer its impact onto the next generations was explained by the idea of heredity. While the notion that morality was inheritable was widely accepted, the question remained how it was to be implemented on a daily basis. Various well-known researchers assigned this function to the brain by associating it with morality. Baldwin (1899) took the brain as a precondition for morality. He suggested that the brain was necessary for a normal development of morality in humans. However, he referred to it in a rather imprecise fashion. He claimed that it must be available, but in his writing its actual function and functioning remained nebulous. Von Krafft-Ebing (1900) claimed that a normal brain is a prerequisite for moral behavior. But he also was unspecific on how the brain is supposed to control morality. Bleuler identified a part of the brain as the seat of morality, though he rejected the idea of a separate moral organ in the phrenological sense. He posited that “inborn or acquired inferiorities” (Bleuler, 1896 p. 21) of the cortex could lead to defective morality. Bleuler did not
think that the seat of morality was confined to a clearly definable area. His identification of the cortex as the neurobiological substrate for morality was vague. Interestingly, Bleuler thought the moral capacity could be disturbed in isolation, if the assumed cortical areas for morality were disrupted. This reasoning is imbued with the prominent theory of brain functioning of the time: localization theory (Hagner, 2008). Specific brain areas were taken to have specific functions, exclusively. In this reasoning, the brain is compartmentalized. Accordingly, the search for a cerebral manifestation of morality was confined to defined regions (Verplaatse, 2009).

Furthermore, the relation of the brain to morality was not specified in detail at the end of the 19th century. This imprecision reflects the fact that researchers were just starting to investigate the brain on a large scale (Finger, 2005). Knowledge about the brain was still meager. The available methods and technologies limited the research options. The research of the day often involved operating on the brain; it was not the present neuroscience with its non-invasive imaging methods. For instance, Ferrier and Jackson were experimenting with induced lesions in animals to localize the motor cortex, but morality could not be examined in a similar vein. Jackson conjectured a hierarchical organization of the brain in which the evolutionarily younger, more advanced prefrontal cortex controlled the less developed regions and thus the socially less desirable urges of a person (Draaisma, 2009). However, scientists’ access to the moral brain was limited. Another way to get to the brain was needed. Hence, at the end of the 19th century the observation of behavior in the morally disturbed was common. Von Krafft-Ebing (1900), for example, documented numerous cases of moral transgressors; sometimes he alluded to their deficient brains. Yet, ideas on a cerebral cause for moral insanity or moral imbecility were of a speculative nature (E. Müller, 1899). Irrespective of the lack of empirical access, the brain was taken to be a necessary precondition for morality. However, it was not a sufficient condition. Under the influence of Darwinian theory, heredity was considered to be
the major force to guarantee for a moral person; the brain was merely associated. Furthermore, it had to share its featured part with the environment, as social forces were also thought to affect the moral person.

It is important to note that inquiries into the moral brain were not mainstream research at the time. Attempts to localize the moral center in the brain were sporadic and proponents of it often had to face fierce criticism (Verplaatse, 2004). Abstractly speaking, morality was not located in the brain on a large scale. Moreover, brain science in general was still in its infancy. The cutting-edge research methods and technologies of the time were laborious. Empirical evidence was scarce and the science itself was only at the beginning of being institutionalized (Star, 1989). Correspondingly, the language of the neuroscientific trade was nascent at this point. The rich technical vocabulary that is common to modern neuroscience had not yet been devised. The possibilities to translate morality by rephrasing it with brain-pidgin were limited. At this point in time, the brain and morality did not speak the same language. Furthermore, the experimental doctrine that is now accepted within psychology was not widely established (Danziger, 1990). As a consequence, scientists were only just beginning to define and operationalize their own concepts that they could subsequently study in experiments. Hence, efforts to define morality in terms of the brain in order to test it experimentally were not as common as today.

Accordingly, the association of the cerebral and the moral realm was unspecific: the relation between brain and morality remained loose. Morality remained elusive in terms of the brain. Yet, it was furnished with some biological features: heritability and evolutionary functionality.

The interlude
The scattered brain-focused studies of morality abated around 1930 (Verplaatse, 2009). A biological approach to morality was reintroduced in the form of sociobiology. Sociobiologists seek to explain human social behavior with the aid of
Invoking the brain in studying morality

evolutionary biology. Around the middle of the 20th century, sociobiologists were particularly interested in the phylogeny of morality (Stent, 1978). Hence, morality was understood to be the result of evolutionary pressures and was considered to be molded by the entire evolutionary history of humanity. It retained the biological features assigned to it at the beginning of the century (heritability and evolutionary functionality), but the tone was lowered with respect to their determinative power. Evolution was merely thought to endow humans with a foundation for morality. This general framework is inscribed in and implemented by genes. Building on this foundation – that is, working within the constraints of the framework – the (social) environment and the brain inform moral behavior. The paramount forces of genes and the environment were indisputable at the time. In this reasoning, the brain is the switch-point between these internal and external powers.

Moreover, a more distinct knowledge emerged about the biological precursors of morality (genes and brains) owing to new methods and technologies, such as cybernetic modeling and simulation (Bischof, 1978), giving rise to new forms of thinking and knowing about morality. Advanced knowledge about the functioning of the brain led to more refined theories about its impact on moral mind and behavior. For example, Eysenck suggested that the brain stem was the basis for individual differences in morality. Furthermore, he stated that “genetically determined cortical arousal” (Eysenck, 1976 p. 123) is pivotal (see Rafter, 2008).

In the genetically informed thinking about morality in the 1970s and 1980s, the function of the brain is more defined than at the beginning of the century. It is pre-programmed by genes and constantly formed by the environment. It is the overlapping entity in the interplay between nature and nurture. As such, it was promoted from the rank of associated feature to become an essential tie in the gene–environment network. The brain here is a mediator. Its role is becoming more visible, but it does not play the main part. Nonetheless, the brain–morality relation is
tightened. Morality is pulled towards the brain as its significance becomes more pronounced. However, genes are at the center of attention; they are the potential locus of morality. Morality is housed in genes rather than in neurons. Moreover, in sociobiological reasoning, morality is defined in terms of altruism and egoism (Dawkins, 1976). The multitude of brain-based definitions that is characteristic of contemporary neuroscience is non-existent.

The present surge
An abundance of studies has recently been published on the putative neural bases of more and more complex aspects of human social life. Among them is a rapidly growing set dealing with traces of morality in the brain (Fumagalli & Priori, 2012; Illes, Kirschen, & Gabrieli, 2003). Though basing ethical qualities in the brain has a long tradition (see above), the current surge in research has a new quality to it. Its marked feature is its dependence on technological developments in neuroimaging, especially functional magnetic resonance imaging. As neuroimaging grants access to the brain, brain research accesses complex social phenomena. A shift in attention from the gene to the brain is brought about, although the genetic approach has not disappeared completely (Raine, 2008). The brain now allows for a potential cerebral reification of morality. The moral mode of operation can allegedly be studied online: that is, in the scanner.

The general procedure of the present approaches is to measure and map the brain. Based on this preprocessing, different approaches to determine brain functioning are launched. Aside from identifying regions of the brain, some researchers focus on the influence of neurotransmitters, such as serotonin (Crockett, Clark, Hauser, & Robbins, 2010), or hormones, such as oxytocin (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Moll & Schulkin, 2009). The shared feature here is the specification of some biological substance that evidently influences the brain, but also influences morality. Yet, the bulk of the research utilizes neuroimaging (e.g.
Greene, Sommerville, Nystrom, Darley, & Cohen, 2001) and to a lesser extent lesion studies (S. W. Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; R. J. Dolan, 1999) to identify brain regions associated with morality.

It is noteworthy that the theoretical framework for this undertaking is neither early 19th-century phrenology nor its historical successor, localization theory. The idea that there are exclusive brain areas for specific mind states or behaviors has vanished. It is widely acknowledged that there is no single moral center in the brain. A specific region may be active in various, unrelated situations. The brain’s complexity and connectivity are proclaimed. Hence, today’s leading framework is a sort of network theory that connects distinct regions to each other and emphasizes their interplay (Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005; L. Young & Dungan, 2012). The same brain region can have more than one function, depending on what the status of the network is. Correspondingly, the brain is now conceived to be a non-centralized, highly interrelated, and dynamic compound. The modern neuroscience of morality tries to make sense of this by identifying patterns of neural activity in connected regions. As a result, the search for the moral center is relinquished and replaced by the search for the “moral neural circuit” (Glenn, Raine, & Schug, 2009 p. 5).

In summary, the introduction of neuroimaging technology in combination with the conviction that the brain is even more complex than previously imagined has made things more convenient and more complicated at the same time. Yet, the brain has become the major unit of analysis because it is widely considered a most fertile ground for researching human mind and behavior, including morality (Rafter, 2008). Along with this came a promotion in power: The brain’s alleged causal influence on morality has never been greater. Although the majority of the neuroscientists concede that culture plays a role in the formation of morality, the brain is now a sufficient condition for morality (Casebeer, 2003). Despite being putatively locked tightly, the
brain-morality relation retains some leeway owing to the usage of imprecise verbs to connect the neural and the moral domains. Specific areas “play important roles” (Prehn et al., 2008 p. 33; Takahashi et al., 2008 p. 1886), “underlie” moral processes, or “are involved in complex social and moral reasoning” (M. B. Miller et al., 2010 p. 2215). As the needed bridge principle between brain and mind is still contested, researchers appear to shy away from making explicit reference to it.

This brief sketch of the history of the relation between the brain and morality is by no means all encompassing. It serves to show how time-specific thinking about morality is deeply entangled in the major theories on neuro-/biological functioning of a given period. Whereas the brain has been loosely associated with morality at the end of the 19th century, it declined in significance until its resurrection as a gene–environment mediator as a consequence of the rise of sociobiology in the middle of the 20th century. In the last two decades, the brain was moved to the center of attention and has become the primary unit of analysis for morality within a certain branch of neuroscience. In summary, the brain has ascended from an associated feature to a potent causal force. It has become the moral domain. The invocation of the brain in studying morality has presently reached its hitherto peak. Moreover, the historical sketch demonstrates that what is being invoked in the study of morality is a dynamic set of notions, not a constant material entity. Thus, the notion of the brain and its relation to morality is mutable. As a consequence, what neuroscientists take morality to be is contingent to a time and place.

A new concept emerges: Brain-based morality

One of the most pronounced characteristics of the modern neuroscience of morality is the explanatory power and informative value assigned to the brain. Understanding the cerebral realm is deemed to be sufficient to understand the moral
realm. This reasoning presupposes that morality is installed in the brain and can be explored in neuroscientific terms. Some labor on the conceptual level is necessary to make this reasoning intelligible and the associated research approach generative. The triadic process of locating, translating, and defining introduced to describe the constitution of gluttony as a neuroscientific object illustrates the required conceptual transformation. To begin with, the ontological status of morality is altered by locating it in convolutions, brain cells, and neurotransmitters. Not only is morality as a whole embedded in the brain, but its diverse aspects are meticulously replaced with certain areas: moral cognition is situated in prefrontal regions, moral emotions can be found in the limbic system, and the social dimension of morality resides in the temporal lobe (Moll et al., 2005). The good and the bad are anchored in the skull at multiple spots. Furthermore, morality’s new locality allows for translating it into the elaborate language of modern neuroscience. For example, Shenhav and Greene (Shenhav & Greene, 2010 p. 670) “found that increased utilitarian responding at the individual level was correlated with increased BOLD activity in bilateral regions of lateral OFC/vlPFC and medial superior frontal gyrus, left middle temporal gyrus and superior parietal lobe.” The new moral semantic comprises neural states and regions. Connected with this is a novel opportunity to learn about morality. Neural activation is correlated with moral engagement, thus scanning people’s brains allows for an educated guess regarding their psychological functioning as well as their virtues and vices. While this way of reasoning is an intended overstatement of neuroscience’s capabilities, it signifies the shift towards the brain as an adequate domain to obtain knowledge about right and wrong. Defining morality in this neuroscientific fashion allows the brain to be a novel epistemological gateway. Consequently, the merging of the cerebral and the moral sphere has brought about a new place of residence, a new language, and new possibilities of knowledge for morality. These modifications have created a novel object of research: brain-based morality. This is a subset of the larger
notion of morality that has been advocated in the history of studying morality without the aid of the brain (e.g. Kurtines & Gewirtz, 1991).

The key aspect in the generation of brain-based morality is the conversion of “knowledge from one domain to another” (Nersessian, 1992 p. 19); this means that neuroscientific knowledge is transferred into knowledge about morality. This refers to a transfer of features. The current notion of the brain imposes a set of features – biological, empirically accessible, sub-personal, network/modular functioning – on morality. As a consequence, the neuroscience of morality views its object of interest as biologically determined and comprised of neurobiological attributes. The transfer of knowledge consists in considering the brain’s mode of operation to be morality’s mode of operation. The research approach also conceives of morality to be measurable in a natural science fashion by experimentally analyzing its material components. Accordingly, the applied methods and technologies aim at capturing somatic properties. Furthermore, morality is deemed to reside in a subsurface domain that is unreachable, and, thus, uncontrollable for the person. Whereas the person is relieved of her or his accountability, the brain is furnished with responsibility. The responsive, conscious agent is thus rendered an illusion and the cerebral subject is endorsed. Lastly, as the current theory of brain functioning implies cerebral modules that interact in a network fashion, this modular structure is installed in morality (see Uttal, 2001, on modularity of mind and brain). Koenigs et al. (2007) state that patients with lesions to the ventromedial prefrontal cortex (“a brain region necessary for the normal generation of emotions and, in particular, social emotions,” p. 908) exhibit a specific tendency in their moral judgments. The brain can be partly destroyed, yet the remaining parts uphold functioning. The statement that fractional brain deficits affect parts of moral functioning indicates a transfer of this modular reasoning. The same approach is applied when moral reasoning is associated with the prefrontal cortex (Casebeer & Churchland, 2003), when it is purported that “the amygdala is important
for affective responsiveness to moral transgressions” (Berthoz, Grèzes, Armony, Passingham, & Dolan, 2006 p. 945), or when “distinct brain areas [are] involved in different kinds of guilt processing” (Basile et al., 2011 p. 238). Morality is split into modules and certain regions of the brain function as stand-ins for certain aspect of morality. In all cases, the functioning of the brain is taken to be the functioning of morality. In this view, morality is not a self-contained entity, but a modular compound, just as the brain.

It is important to note that morality poses some constraints on the transfer of attributes. The most evident constraint is the cultural versatility of morality; therefore, strict brain determinism is kept at bay. Neuroscientists do not maintain that because we all have the same brains, we all have the same morality. While the content of morality remains indefinite in terms of the brain, the functioning of morality is taken to be universally determined by it (Haidt & Joseph, 2004). In summary, the transfer of knowledge from the biological to the moral realm results in a change of meaning for morality. To certain neuroscientists, morality now has a cerebral connotation.

Some qualifications are necessary prior to concluding. Although the study of brain-based morality is popular within a certain scientific niche as well as the media (O'Connor, Rees, & Joffe, 2012), the practice is by no means unanimously welcomed. Neuroscientists who endorse brain-based morality face controversy and criticism from various angles. Inside the field, there is an ongoing debate on the functioning of the moral brain (e.g. L. Young & Dungan, 2012). From outside the discipline, the neuroscience of morality is challenged from advocates of behavioral moral psychology (Turiel, 2008), indicted for methodological confusion (Kahane & Shackel, 2010; Robinson, 2012; Schleim, 2008), critically assessed from legal scholars (Morse, 2006), and contested from an ethical perspective (Illes, Sahakian, & Federico, 2011).
Chapter 1

**Conclusion**

The invocation of the brain in studying morality is multifarious. It is a popular research field that receives growing attention. Furthermore, it is a development within the wider context of the neurobiologization of psychology. Also, it is a historical process which entails versatile and contingent notions of the brain and morality. Lastly, and most importantly here, it comprises the establishment of a new concept of morality. In order for neuroscience to study morality, it needed to adjust its meaning to create a viable target for itself. Morality needed to be transformed in order to be handled by neuroscientific theories, methods, and machinery. As a consequence, the ontological, epistemological, and semantic attributes of morality are being located, translated, and defined in terms of the brain. Correspondingly, morality now exists in the brain, it is sought there, and it is talked about in terms of it. By attaching morality to the brain, neuroscientists restricted the elusiveness and dynamicity of morality; morality as an object of knowledge is downsized into brain-based morality. The study of this specific object of research resulted in the establishment of a self-contained research field with its own body of knowledge. The key point is that the invocation of the brain in studying morality was not a neutral act: the neuroscience of morality is altering what it seeks to study.