

# What is neuroethics? Empirical and theoretical neuroethics

Georg Northhoff

Canada Research Chair for Mind, Brain and Neuroethics, Michael Smith Chair for Neuroscience and Mental Health, Institute of Mental Health Research, University of Ottawa, Ottawa, Ontario, Canada

Correspondence to Georg Northhoff, Institute of Mental Health Research University of Ottawa, 1145 Carling Avenue, Ottawa, ON K1Z 7K4, Canada  
Tel: +1 613 722 6521; fax: +1 613 798 2982;  
e-mail: Georg.northhoff@rohcg.ca

**Current Opinion in Psychiatry** 2009, 22:000–000

## Purpose of review

Neuroethics is a recently emerging field that deals with predominantly empirical and practical issues of ethics in neuroscience. In contrast, theoretical and methodological considerations have rather been neglected and thus what may be called theoretical neuroethics.

## Recent findings

The review focuses on informed consent and moral judgment as examples of empirical neuroethics and norm–fact circularity and method-based neuroethics as issues of a theoretical neuroethics.

## Summary

It is argued that we need to consider theoretical and methodological issues in order to develop neuroethics as a distinct discipline, which as such can be distinguished from both philosophy/ethics and neuroscience.

## Keywords

informed consent, method-based neuroethics, moral judgment, norm–fact circularity

Curr Opin Psychiatry 22:000–000  
© 2009 Wolters Kluwer Health | Lippincott Williams & Wilkins  
0951-7367

## Introduction

The recent progress in neuroscience has led to ethical questions concerning the emergence of a novel field, neuroethics. Neuroethics can broadly and preliminarily be defined as drawing relationships between neuroscientific observations and ethical concepts. Roskies [1] distinguished between ethics of neuroscience and neuroscience of ethics. The ethics of neuroscience deal with ethical problems in neuroscience arising from new forms of interventions into the brain, whereas the neuroscience of ethics investigates the neural mechanisms that may possibly underlie ethical concepts and practices like informed consent, moral judgment, free will, and so on.

No sharp distinction exists between the neuroscience of ethics and the ethics of neuroscience. Consider the example of informed consent: How valid is the informed consent of patients whose cognitive and emotional capacities are altered or deteriorated by the nature of their disorder, as, for instance, in dementia or schizophrenia? This is an issue in the domain of ethics of neuroscience. It has also triggered empirical investigations of those cognitive and neural functions that are implicated in giving a valid informed consent that falls more into the domain of neuroscience of ethics (see [2] as well as [3<sup>••</sup>]).

Another such example where neuroscience of ethics and ethics and neuroscience converge is the recent discussion

about moral judgment. What is a moral judgment and how does it affect our ethical decisions in the current neuroscience? The question about the nature of moral judgment has triggered many neuroscientific investigations of the neural mechanisms underlying moral judgment (see [4]) that touches upon the neuroscience of ethics. However, the focus on the impact of moral judgment on our ethical decisions in neuroscience may rather be considered a question for the ethics of neuroscience. The discussion of moral judgment thus provides one possible node point where both neuroscience of ethics and ethics of neuroscience intersect and converge.

Although a clear-cut distinction between neuroscience of ethics and ethics of neuroscience seems to remain unclear, they share their focus on empirical issues; be they predominantly neuroscientific, as in the neuroscience of ethics, or rather practical or ethical, as in the ethics of neuroscience. One may consequently subsume both ethics of neuroscience and neuroscience of ethics under the umbrella what may be called empirical neuroethics. The first aim of my paper is to highlight some recent development in empirical neuroethics, thereby focusing on informed consent and moral judgment.

## Empirical neuroethics

Empirical neuroethics focuses on the empirical, that is, psychological and neural conditions that may underlie ethical concepts like informed consent, free will, and so on.

## 2 History and philosophy

### Informed consent and emotions

Informed consent is crucial in clinical scientific studies of healthy individuals and psychiatric patients. This has led to an extensive research of the factors and functions determining the kind of consent psychiatric patients give, that is, whether they accept, refuse, or remain ambivalent about study participation (see [2,5–13]). Giving informed consent requires the capacity to make a decision, that is, decision making [7,8,12–20]. Decision making is a complex process that involves both cognitive and affective functions, as pointed out by recent neuroscientific research [2,21,22].

Cognitive functions in decision making concern attention, working memory, executive functions, and others (see [7,8,13–20,23]). More specifically, the capacity to make a decision in informed consent, the so-called decisional capacity, involves cognitive functions like understanding, appreciation, and reasoning. These cognitive functions and, therefore, decisional capacity have recently been systematically investigated with the MacArthur Competence Assessment Tool for Clinical Research (MacCAT-CR) [5,11]. The MacCAT-CR focuses predominantly on cognitive functions like understanding, appreciation, reasoning, and expressing a choice that are supposedly implicated in the decisional capacity for informed consent.

In addition to cognitive functions, decision making in general and decisional capacity in informed consent in particular involve empathy and emotions [2,6,9,10,21,22,24]. Empathy describes the ability to share another person's cognitive and emotional inner life (see [25,26,27,28,29,30,31]). Empathic sharing between the consenting individual and the investigator may be crucial in determining decisional capacity and ultimately the consent itself. This is possible only when the consenting person is able to recognize the investigator's emotions, thus requiring emotion recognition [32]. In contrast to cognitive functions, the impact of empathy and emotion recognition on decisional capacity and the consent process has yet to be investigated empirically.

The involvement of specific cognitive and affective functions in informed consent may also provide some clues about its possible underlying neural mechanisms. Cognitive functions like working memory and attention are well known to be associated with neural activity in the lateral prefrontal cortex, whereas emotional functions rather involve medial cortical and subcortical regions [33,34]. If so, the balance between, for instance, medial and lateral prefrontal cortical functions may be crucial in constituting the kind of decision making that is crucial in informed consent. Interestingly, recent imaging studies demonstrated reciprocal modulation between neural activity in medial and lateral prefrontal cortex during

affective and cognitive functions [34]. This suggests that a valid informed consent may be necessarily (though not sufficiently) dependent on a specific balance between affective and cognitive functions and thus between medial and lateral prefrontal cortex. If the medial–lateral prefrontal cortical neural activity is altered and unbalanced, as, for instance, in depression and schizophrenia (see [35,36,37]), the constitution of decision making implicated in informed consent may be affected. This in turn may diminish the validity of the consent. Such scenario remains highly speculative though at this point because no imaging study during the decision making of informed consent has yet been conducted.

### Moral judgment

We make moral judgments daily like whether we should help the person in front of us to get up to his feet even if it means that we would then come too late to work, entailing possible rebuff by the boss. These kinds of situations present us with a moral dilemma whether we should act in the interest of others or rather according to our own goals and interests. Various imaging studies [4,38,39,40] presented sentences where persons described moral dilemma situations where they had to choose between self-interest and moral decisions, that is, decisions that were in favor of other persons rather than their own. These imaging studies revealed the regions that are particularly active during moral action and judgment. Various studies (see [4,38,39,40]) demonstrated the involvement of the medial prefrontal cortex (MPFC) and the ventral striatum in moral judgment.

This, however, raises several problems. The above-mentioned regions, the MPFC and the ventral striatum, are not only involved in moral judgment but also in various other processes ranging from emotion, mentalizing with inference of other's mental states, learning of reward contingencies (reward representation), and thinking about one's own mental states resulting in self-awareness (see [41] for an overview). This raises not only the question about the anatomical specificity of these regions for moral judgment but also about the psychological components implicated in moral judgment.

Does moral judgment implicate various psychological functions as diverse as learning, reward, mental state attribution, theory of mind, and so on? How can we characterize moral judgment in both psychological and neural terms? The above-mentioned studies clearly indicate that emotions and empathy are central psychological processes in moral judgment (see also [42]). What does this imply in ethical regard? Does this mean that people with deficits in emotions and empathy show reduced moral judgment? Psychopaths show indeed severe deficits in their own emotions and the sharing and understanding of other's emotions, empathy that also

seems to go along with deficits in medial prefrontal cortical activation. Does this mean that they are unable to make appropriate moral judgments? In the case of psychopaths, this is certainly true at a behavioral level, as they do indeed often commit crimes, thereby remaining unaffected by both own and others' emotions. One would consequently hypothesize that psychopaths who committed crimes may show severe deficits in medial prefrontal cortical activity during moral judgment. This remains to be tested though in future empirical studies.

In contrast to the empirical issues, theoretical issues have largely been neglected in current neuroethics. How does the kind of ethical concept, as presupposed in neuroethics, impact both practical issues of its application in neuroscience and study designs for the investigation of its neural correlates? And how do the empirical findings in both ethics of neuroscience and neuroscience of ethics impact the definition of the ethical concept in question? This raises also a more methodological rather than a conceptual issue. How can we make the translation between ethical concepts and neuroscientific findings? We do need predefined rules and valid methodological strategies for linking ethical concepts and neuroscientific findings. These conceptual and methodological issues may be subsumed under the umbrella of theoretical neuroethics. The second aim is to provide a brief discussion of conceptual and methodological issues that fall into the domain of theoretical neuroethics.

---

### **Theoretical neuroethics**

Empirical neuroethics deals with the empirical and practical aspects of the linkage between neuroscientific and ethical concepts. Theoretical neuroethics, in contrast, focuses on the methodological and conceptual aspects of such linkage that allow to link ethical concepts and neuroscientific facts and thus descriptive and normative dimensions. Although there has been much discussion of various issues in empirical neuroethics, the discussion of methodological and conceptual issues and thus theoretical neuroethics has remained rather sparse so far. Owing to the vital importance of these issues for the future mapping of neuroethics as a distinct and separate discipline in its own right, I want to devote the second part of this review to a brief account of theoretical neuroethics.

### **Methodological confusions between norms and facts**

Are moral judgments 'located' in the above-mentioned regions? Is moral judgment nothing but the neural activity in these regions? This is no longer an empirical question but one that touches upon conceptual and methodological issues that fall into the domain of what I call theoretical neuroethics.

Conceptually, neuroethical concepts like moral judgment and informed consent must be hybrids, conceptual hybrids as I call them. A conceptual hybrid denotes a single concept that contains or implies two or more different types of concepts, as for instance normative and descriptive concepts. Neuroethical concepts may be characterized as conceptual hybrid, as they are neither purely normative concepts, as for instance concepts from philosophical ethics, nor purely descriptive, as concepts in neuroscience. The hybrid nature of neuroethical concepts thus consists in the linkage between normative and descriptive dimensions, that is, between norms and facts. One may go even one step further and argue that this linkage between norms and facts defines neuroethical concepts as neuroethical. If there is no such norm–fact linkage, neuroethical concepts would degenerate either into ethical and thus purely normative concepts or neuroscientific concepts as seemingly purely descriptive.

To identify moral judgment with activity in specific brain regions is thus to commit two confusions. First, one confuses the necessary conditions, that is, the neural mechanisms, and what they condition, that is, the moral judgment as the result. This also implies that one may neglect the difference between necessary and sufficient conditions and to regard the neural facts not only as necessary but also as sufficient condition.

Second, one also confuses descriptive and normative components of moral judgment. The neural observations describe facts, the descriptive component, whereas the moral judgment also implicates a normative component that goes beyond the descriptive component by referring to normative dimensions. If one now infers what kind of brain state is morally correct from the empirical findings about moral judgment, one confuses descriptive and normative components and thus facts and norms. This is problematic though, as one cannot infer the latter, that is, norms, from the former, that is, facts, which would mean to neglect their principal difference.

### **Linkage between norms and facts in neuroethical concepts**

The question is now how we can link the two different dimensions, norms and facts, within the neuroethical concepts. We have different choices. One may disregard the normative level and reduce it to the factual one; ethical norms are then unilaterally replaced by neural facts, that is, unilateral replacement. This, however, is to neglect the principal difference between norms and facts. Alternatively, one may accept the norm–fact distinction and consider norms and facts in a parallel or bilateral way, resulting in what may be called bilateral parallelism.

However, both solutions, unilateral replacement and bilateral parallelism, do not do justice to the close

#### 4 History and philosophy

intertwining of ethical concepts and neuroscientific observations in the current neuroethical debate. Unilateral replacement of norms in favor of facts falls short of considering the idiosyncratic nature of the normative component when compared to the descriptive one. As pointed out in the second confusion, we cannot infer norms from facts. However, bilateral parallelism falls short of taking into account the close interdependence between normative and descriptive components and thus between norms and facts. The current neuroethical debate in general and the above-described examples of moral judgment and informed consent in particular point out that changes in one, as for instance the descriptive component, may entail changes in the other, the normative components.

What we need is a methodological strategy to link norms and facts, that is, ethical concepts and neuroscientific findings, in a much closer and tighter way without falling into either unilateral replacement or bilateral parallelism. We may need to devise systematic methodological strategies to link norms and facts in neuroethical concepts that account for their principal difference as well as for their close interdependency.

One such methodological strategy may be to go back and forth, that is, to circulate or oscillate between ethical concepts and neuroscientific findings and thus between norms and facts; I call this ‘norm–fact circularity’. How can we briefly define such norm–fact circularity? The usual starting point of empirical neuroethics is an ethical concept that is linked to neuroscientific observations. The aim here is to either ‘neuronize’ the ethical concept, as neuroscience of ethics, or to reveal its relevance in neuroscientific research, as in ethics of neuroscience. What, however, is often neglected is how this first encounter of ethical concepts with neuroscientific observations bears upon the ethical concept itself. What is implied for the concept of informed consent if informed consent is driven empirically by emotions and empathy rather than cognitive functions? What does the fact of emotion/empathy imply for the norms inherent in informed consent? Do we have to conceptualize the linkage between norms and facts in informed consent (and moral judgment) in a different way if emotions/empathy are predominant rather than cognitive functions?

We may thus make conceptual modifications in neuroethical concepts depending on the neuroscientific findings. This then makes the initial ethical concept a truly neuroethical concept in a literal rather than merely figurative sense. The circularity between ethical concepts and neuroscientific findings goes, however, one step further. The revised and modified neuroethical concepts may make different empirical approaches and

study designs in subsequent neuroscientific investigations necessary in order to raise further empirical support. There is such circularity or a loop between ethical concepts and neuroscientific findings and thus between norms and facts – I call this ‘norm–fact circularity’.

#### Method-based neuroethics

The concept of systematic norm–fact circularity describes a methodological strategy to account for the close interdependence between norms and facts without neglecting their principal differences. As such, it aims to describe how normative and descriptive dimensions in neuroethical concepts can be linked to each other in a systematic way that goes beyond mere intuitive connection as in bilateral parallelism or unilateral replacement. What is needed in the future is the development of a systematic method that details the different steps norm–fact circularity by giving exact methodological prescriptions and measures of validity and reliability.

One may now argue that such ‘systematic norm–fact circularity’ is a merely theoretical playground without any implications for empirical neuroethics. This, however, is to neglect the need for a specific neuroethical methodology. If neuroethics wants to establish itself as a separate discipline that is different from its neighboring disciplines like philosophy, ethics, and neuroscience, it must develop a special methodology.

Only the development of a specific methodology (or even several methodological strategies) will allow neuroethics to establish itself as a discipline with its own rights as different from philosophy/ethics and neuroscience. As such, it needs to be based more on a (or seven several) specific methodological strategy(ies) rather than a certain result, as for instance in unilateral replacement with the consecutive neuronization of ethical concepts. One may consequently speak of a method-based neuroethics as distinguished from result-based neuroethics.

---

#### Conclusion

I here gave a brief review of the distinct aspects or dimensions being prevalent in current neuroethics. These included empirical aspects that focus on the practical and empirical issues arising from the encounter between ethical concepts like moral judgment and informed consent and neuroscientific findings as in emotions and empathy. Such neuroethical encounter presupposes the principal possibility of linking normative and descriptive dimensions and thus norms and facts. The second part of the review, therefore, focused on some often neglected theoretical and methodological issues on how facts and norms can be linked in an interdependent way without neglecting their principal difference. This makes the elaboration of systematic

AQ1

methodological strategies as specific neuroethical necessary, which may then provide the ground for the future development of a truly method-based neuroethics.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 000–000).

- 1 Roskies A. Neuroethics for the new millennium. *Neuron* 2002; 35:21–23.
- 2 Northoff G. Neuroscience of decision making and informed consent: an investigation in neuroethics. *J Med Ethics* 2006; 32:70–73.
- 3 Levy N, Clarke S. Neuroethics and psychiatry. *Curr Opin Psychiatry* 2008; •• 21:568–571.  
This paper gives an excellent overview about ethical issues in psychiatry.
- 4 Greene JD, Sommerville RB, Nystrom LE, *et al.* An fMRI investigation of emotional engagement in moral judgment. *Science* 2001; 293:2105–2108.
- 5 Appelbaum PS, Grisso T, Frank E, *et al.* Competence of depressed patients for consent to research. *Am J Psychiatry* 1999; 156:1380–1384.
- 6 Breiden TM, Vollmann J. The cognitive based approach of capacity assessment in psychiatry: a philosophical critique of the MacCAT-T. *Healthcare Anal* 2004; 12:273–283; discussion 65–72.
- 7 Candilis PJ, Geppert CM, Fletcher KE, *et al.* Willingness of subjects with thought disorder to participate in research. *Schizophr Bull* 2006; 32:159–165.
- 8 Carpenter WT Jr, Gold JM, Lahti AC, *et al.* Decisional capacity for informed consent in schizophrenia research. *Arch Gen Psychiatry* 2000; 57:533–538.
- 9 Charland LC. Appreciation and emotion: theoretical reflections on the MacArthur Treatment Competence Study. *Kennedy Inst Ethics J* 1998; 8: 359–376.
- 10 Charland LC. Affective neuroscience and addiction. *Am J Bioeth* 2007; 7:20–21.
- 11 Grisso T, Appelbaum PS, Hill-Fotouhi C. The MacCAT-T: a clinical tool to assess patients' capacities to make treatment decisions. *Psychiatr Serv* 1997; 48:1415–1419.
- 12 Jeste DV, Palmer BW, Golshan S, *et al.* Multimedia consent for research in people with schizophrenia and normal subjects: a randomized controlled trial. *Schizophr Bull* 2008. [Epub ahead of print]
- 13 Koren D, Poyurovsky M, Seidman LJ, *et al.* The neuropsychological basis of competence to consent in first-episode schizophrenia: a pilot metacognitive study. *Biol Psychiatry* 2005; 57:609–616.
- 14 Candilis PJ, Fletcher KE, Geppert CM, *et al.* A direct comparison of research decision-making capacity: schizophrenia/schizoaffective, medically ill, and nonill subjects. *Schizophr Res* 2008; 99 (1–3):350–358.
- 15 Howe V, Foister K, Jenkins K, *et al.* Competence to give informed consent in acute psychosis is associated with symptoms rather than diagnosis. *Schizophr Res* 2005; 77 (2–3):211–214.
- 16 Moser DJ, Schultz SK, Arndt S, *et al.* Capacity to provide informed consent for participation in schizophrenia and HIV research. *Am J Psychiatry* 2002; 159:1201–1207.
- 17 Palmer BW, Dunn LB, Appelbaum PS, Jeste DV. Correlates of treatment-related decision-making capacity among middle-aged and older patients with schizophrenia. *Arch Gen Psychiatry* 2004; 61:230–236.
- 18 Palmer BW, Jeste DV. Relationship of individual cognitive abilities to specific components of decisional capacity among middle-aged and older patients with schizophrenia. *Schizophr Bull* 2006; 32:98–106.
- 19 Roberts LW, Warner TD, Brody JL, *et al.* Patient and psychiatrist ratings of hypothetical schizophrenia research protocols: assessment of harm potential and factors influencing participation decisions. *Am J Psychiatry* 2002; 159: 573–584.

AQ2

- 20 Stroup S, Appelbaum P, Swartz M, *et al.* Decision-making capacity for research participation among individuals in the CATIE schizophrenia trial. *Schizophr Res* 2005; 80:1–8.
- 21 Damasio AR. *The feeling of what happens: body and emotion in the making of consciousness.* New York: Harcourt Brace; 1999.
- 22 Northoff G, Grimm S, Boeker H, *et al.* Affective judgment and beneficial decision making: ventromedial prefrontal activity correlates with performance in the Iowa Gambling Task. *Hum Brain Mapp* 2006; 27:572–587.
- 23 Jeste DV, Depp CA, Palmer BW. Magnitude of impairment in decisional capacity in people with schizophrenia compared to normal subjects: an overview. *Schizophr Bull* 2006; 32:121–128.
- 24 Kluge EH. Competence, capacity, and informed consent: beyond the cognitive-competence model. *Can J Aging* 2005; 24:295–304.
- 25 de Vignemont F, Singer T. The empathic brain: how, when and why? *Trends Cogn Sci* 2006; 10:435–441.
- 26 Decety J, Batson CD. Social neuroscience approaches to interpersonal sensitivity. *Soc Neurosci* 2007; 2 (3–4):151–157.
- 27 Dziobek I, Rogers K, Fleck S, *et al.* Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *J Autism Dev Disord* 2008; 38:464–473.  
This paper gives a good introduction into the testing of empathy.
- 28 Fan Y, deGreeck M, Duncan N, Northoff G. Is there a core process in empathy? An fMRI-based metaanalysis of domain-independent regions, in press.
- 29 Hein G, Singer T. I feel how you feel but not always: the empathic brain and its modulation. *Curr Opin Neurobiol* 2008; 18:153–158.  
This is a wonderful review about the neural correlates of empathy, the concepts and the empirical findings.
- 30 Lamm C, Decety J. Is the extrastriate body area (EBA) sensitive to the perception of pain in others? *Cereb Cortex* 2008; 18:2369–2373.
- 31 Singer T, Seymour B, O'Doherty JP, *et al.* Empathic neural responses are modulated by the perceived fairness of others. *Nature* 2006; 439:466–469.
- 32 Bowers D, Blonder L, Heilman K. *The Florida Affect Battery.* Center of Neuropsychological Studies; 1992.
- 33 Grimm S, Schmidt CF, Bermpohl F, *et al.* Segregated neural representation of distinct emotion dimensions in the prefrontal cortex—an fMRI study. *Neuroimage* 2006; 30:325–340.
- 34 Northoff G, Heinzel A, Bermpohl F, *et al.* Reciprocal modulation and attenuation in the prefrontal cortex: an fMRI study on emotional–cognitive interaction. *Hum Brain Mapp* 2004; 21:202–212.
- 35 Grimm S, Beck J, Schuepbach D, *et al.* Imbalance between left and right dorsolateral prefrontal cortex in major depression is linked to negative emotional judgment: an fMRI study in severe major depressive disorder. *Biol Psychiatry* 2008; 63:369–376.
- 36 Grimm S, Boesiger P, Beck J, *et al.* Altered negative BOLD responses in the default-mode network during emotion processing in depressed subjects. *Neuropsychopharmacology* 2009; 34:932–943.  
This is a study about depression and its neural network changes suggesting increased hyperactivity in the resting state in depression.
- 37 Grimm S, Schroger E. The processing of frequency deviations within sounds: evidence for the predictive nature of the Mismatch Negativity (MMN) system. *Restor Neurol Neurosci* 2007; 25 (3–4):241–249.
- 38 Moll J, de Oliveira-Souza R. Moral judgments, emotions and the utilitarian brain. *Trends Cogn Sci* 2007; 11:319–321.
- 39 Moll J, de Oliveira-Souza R, Bramati IE, Grafman J. Functional networks in emotional moral and nonmoral social judgments. *Neuroimage* 2002; 16 (3 Pt 1): 696–703.
- 40 Young L, Saxe R. The neural basis of belief encoding and integration in moral judgment. *Neuroimage* 2008; 40:1912–1920.  
A cognitive approach to moral judgment with an excellent study design.
- 41 Legrand D, Ruby P. What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol Rev* 2009; 116:252–282.
- 42 Panksepp J. *Affective neuroscience: the foundations of human and animal emotions.* New York: Oxford University Press; 1998.

AQ3

Dear Author,

During the preparation of your manuscript for typesetting, some queries have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin as neatly as possible or compile them as a separate list. This form should then be returned with your marked proof/list of corrections to the Production Editor.

**QUERIES: to be answered by AUTHOR/EDITOR**

QUERY NO.	QUERY DETAILS	RESPONSE
<AQ1>	Please check whether the intended meaning of the sentence is retained after making the edits.	
<AQ2>	Please update Ref. [12].	
<AQ3>	Please provide the complete bibliographic details in Ref. [28].	