Norms and the Brain

an Investigation Into the Neuroscience of Ethical Decisions

and the Ethics of Neuroscience

Dissertation
zur Erlangung des Doktorgrades (Ph.D.)
des Fachbereichs Humanwissenschaften
der Universität Osnabrück

vorgelegt
von
Stephan Schleim
aus
Wiesbaden

Osnabrück, 2009
Norms and the Brain

an Investigation Into the Neuroscience of Ethical Decisions and the Ethics of Neuroscience

an interdisciplinary dissertation written by Stephan Schleim, M.A.
born on March 8 1980 in Wiesbaden
submitted to the FB 8 Humanwissenschaften at the University of Osnabrück

this publication is based on the dissertation submitted under the same title and defended on December 17 2009 in Osnabrück
Norms and the Brain

*an Investigation Into*

*the Neuroscience of Ethical Decisions*

*and the Ethics of Neuroscience*

## Table of Contents

| Synopsis | Norms and the Brain | 1 |
| Study 1 | From moral to legal judgment: The influence of normative context in lawyers and other academics (*publication note*) | 27 |
| Study 2 | Judging a Normative Judgment: An Exploration into Higher-Order Normative Cognition (*extended abstract*) | 28 |
| Study 3 | Moral Physiology, Its Limitations and Philosophical Implications (*publication note*) | 34 |
| Study 4 | The risk that neurogenetic approaches may inflate the psychiatric concept of disease and how to cope with it (*publication note*) | 35 |
Synopsis

Norms and the Brain

Stephan Schleim
Norms and the Brain

From Neuroscience to Neuroethics

If the 1990s were the “Decade of the Brain”, then the subsequent decade can be called the “Decade of Neuroethics”. “Neuroethics” was introduced as a concept encompassing two different kinds of investigation, first the ethical implications of neuroscience and second the neuroscientific investigation of ethical phenomena, mainly moral perception and decision (Roskies, 2002).

The early 2000s have not only seen major publications within the domain of neuroethics (Glannon, 2007; Illes, 2006; Levy 2007; Merkel et al., 2007; Sinnott-Armstrong 2008), but also the foundation of academic associations (e.g. the Neuroethics Society, the International Neuroethics Network), specialized journals (e.g. Neuroethics), and even specialized institutions and programs at universities (e.g. the Oxford Centre for Neuroethics, the National Core for Neuroethics at the University of British Columbia, the Program in Neuroethics at the Stanford Center for Biomedical Ethics). This development illustrates an increasing sensibility for the wider social, ethical, and legal implications of investigations as well as manipulations of the human brain which are the central topic of neuroethics.

This is not surprising considering the wide spectrum of neuroscientific research projects concerning essentially human capacities such as economic (Glimcher and Rustichini, 2004; Fehr & Camerer, 2007; Krueger, Grafman, and McCabe, 2008), moral (Greene and Haidt, 2002; Moll et al., 2005), or legal decision making (Buckholtz et al., 2008), deception (Sip et al., 2008),

---

1 But see my critique of that notion in Study 3.
2 Of course, that I call them essentially human does not imply that they are uniquely human. Many of the listed capacities have precursors or analogues in the animal kingdom.
empathy (de Vignemont and Singer, 2006; Singer et al., 2006), or romantic love (Bartels and Zeki, 2004) – that is, a subdomain of neuroscience clearly has become social during the last decade (Adolphs, 2003; Lieberman, 2007) and is now addressing questions which exclusively belonged to the domains of psychology, the social sciences, or the humanities before the “Decade of the Brain”.

Even philosophy has been changed by progress in the neurosciences; and this change is not confined to philosophy of science, where philosophers naturally analyze the scientific methods and concepts of neuroscience (Bechtel et al., 2001; Bennett and Hacker, 2004) like others did before for other disciplines. More astonishingly, even traditional philosophical debates on topics such as free will and moral responsibility (Levy, 2006a; Roskies, 2006; Walter, 2001), the role of emotion and intuition in ethics (Joyce, 2008; Levy, 2006b; Singer, 2005), and of course the nature of cognition and consciousness (Block, 2005; Laureys, 2005; Metzinger, 2000) have been strongly influenced by recent empirical research. The present dissertation has to be evaluated against this background, that is, the intermixture of neuroscience, philosophy, and ethics, because “Norms and the Brain” comprises all of these aspects.

The research that is presented in the forthcoming chapters was carried out within and beyond the interdisciplinary “animal emotionale” grant at the Universities of Frankfurt, Bonn, and Osnabrück between 2005 and 2009. Project 3 of this grant concerned the neuroscientific investigation of moral decision making, which had started a few years earlier (Greene et al., 2001;
Moll et al., 2001), as well as the philosophical implications of this research. The intention was not only to probe the influence of emotions on moral judgments, but to analyze whether the new empirical results had any bearing on established philosophical accounts of ethics. In the course of this endeavor, one of the challenges was to bridge both, the empirical and the philosophical frames, in a way that results on the one side could also be evaluated and employed on the other side. It soon became clear that the methodological limitations of functional magnetic resonance imaging, the research tool chosen to investigate moral decisions, constrained the questions that could be addressed in such an experiment; however, it became also clear that the conceptual ambiguity of the term “moral judgment” or, for that matter, even the terms “moral” and “judgment” in isolation, already constrained such attempts independent of any methodological considerations. That is, the project provided challenges for philosophers as well as neuroscientists right from the outset.

Unlike previous research, we decided to investigate moral judgments along with a second kind of normative judgment, namely, legal judgments (Study 1; Study 2). This did not only allow us to see whether these different normative contexts would differ in the engaged emotional and cognitive processes, but also to check the generality of the so-called moral “network” (Moll et al., 2005, p. 799). Our experimental design thus took the problem of the “reverse inference” (Poldrack, 2006) seriously: Because many brain areas are associated with a number of cognitive processes, the presence of the latter cannot simply be inferred from activation of the former. That is, the inferential validity of the reverse inference continuously declines with
an increasing number of different cognitive functions within one particular brain area. Our research thus addressed issues of philosophy of science and theoretical neuroscience which are relevant to research beyond the topic of moral cognition.

These considerations were addressed in more detail and more theoretically in a separate investigation which analyzed the research of moral physiology in general and its philosophical implications in particular (Study 3). Especially, it reviewed the empirical literature with respect to the philosophical claims that had already been inferred from it (Casebeer, 2003; Greene, 2003; Singer, 2005) and suggested conceptual clarifications concerning the domains of moral physiology and neuroethics. I separated descriptive from normative neuroethics in order to avoid a confusion of descriptive and normative issues already on the conceptual level. The former would then comprise empirical research like that performed in Studies 1 and 2 as well as descriptions of ethical problems related to neuroscience; the latter, however, would suggest solutions – ideally, of course, informed by a clear analysis of the respective issue – how to cope with such problems.

It is this field into which Study 4 belongs which investigated the impact that neuroscience and its combination with genetics might have on the understanding and application of the psychiatric concept of disease. Clearly, such an investigation falls into the scope of “Norms and the Brain”, given the normative conceptions and implications of illnesses and the brain-centered view on mental disorders, that is, their increasing understanding in
terms of neural and genetic deficits (Caspi and Moffitt, 2006). This view expresses itself in the diagnosis and treatment of such conditions on the neural level, for example, through neuroimaging approaches, psychopharmacological substances, or even neurosurgical interventions. The normative consequences for individuals diagnosed with psychiatric disorders, their closer social environment, and society as a whole, are vast. It is thus important that patients are diagnosed correctly, treated accordingly, and that sane subjects are not falsely considered as ill. Since the concepts of many psychiatric disorders are ambiguous, fuzzy, and in some cases even highly controversial, I identified the risk and suggested a solution to cope with the possibility that an accordingly prepared experiment based on neuroscientific and genetic methods might be abused to describe normal conditions as pathological – apparently giving this distinction a sound scientific basis – and thus to “inflate” the psychiatric concept of disease.

The four studies comprising this dissertation are thus interconnected through the different steps of analysis and evaluation concerning “Norms and the Brain”: At the outset of this endeavor are Studies 1 and 2 which...
address normative decision making on the very empirical level. They are followed by the more analytical and evaluative Studies 3 and 4 which deal with the philosophical as well as the wider ethical and social implications of such research and neuroscience in general. Ideally, such an analysis and evaluation is in turn followed by further empirical research, analysis and evaluation. The inherent relation of “Norms and the Brain” is also demonstrated in Illustration 1. An outlook on questions that lend themselves to investigation in future research is thus given at the end of this chapter, after the individual summaries of the four studies.
Study 1 – From moral to legal judgment: The influence of normative context in lawyers and other academics

In the first study, we wanted to understand the emotional and cognitive processes involved in moral cognition better, particularly whether emotional influences impact moral judgments. One obvious way to address this issue would have been to influence the emotional content of our stimulus material (e.g. Heekeren et al., 2005) or to instruct subjects explicitly to regulate their emotions (e.g. Harenski and Hamann, 2006). However, we decided against these options and developed a normative judgment task with the following characteristics:

First of all, we acknowledged that there was no clear and uncontested understanding of what a “moral judgment” that would lend itself to empirical research; by contrast, we integrated the task implicitly into our instruction which asked the subjects to make a judgment either from “the moral view of a citizen” or “the legal view of a judge”. That is, we avoided the operationalization of a theoretical construct whose ecological validity would be unclear and decided in favor of the actual understanding our research subjects had. Second, we decided to modify the task instruction, not the stimulus material. In social neuroscience, where the chosen research materials have to be very complex in order to capture essential aspects of the social, it is often difficult to control whether the experimental results are related to the manipulations intended by the researchers or are subject to confounding factors accompanying them (e.g. different emotions, self-relatedness, or autobiographical familiarity are often entangled in such stimulus material). Furthermore, we decided to investigate two different
groups of subjects in our experiment, namely, legal experts and other academics. We thus wanted to find out whether the subject’s experience would have an influence on their cognitive processing during normative judgment and on their decisions.

Nevertheless, our research was not only an exploration to test the specificity of the brain activations associated with moral judgment. Considering the idealistic understanding of legal judgment that is prevalent in the scholarly literature of western societies (Gewirtz, 1996; Goodenough, 2001), we hypothesized that legal decisions would be more subject to influences of cognition than emotion and that this effect would be particularly pronounced in lawyers. We operationalized this assumption according to neuroscientific theories relating the dorsolateral prefrontal cortex (DLPFC) to the application of explicit rules (Bunge, 2004; MacDonald et al., 2000; Miller and Cohen, 2001). By contrast, we expected that moral judgments would be made more intuitively. While our results partially confirmed our hypothesis, the data as a whole did not allow an easy explanation. We indeed found stronger activation within the left DLPFC when subjects made legal judgments as contrasted to moral judgments. This suggests that there is an important difference in the cognitive processing of both kinds of normative judgment. Furthermore, our behavioral data showed that lawyers perceived themselves as emotionally less involved during normative judgment than other academics.

However, the idea that moral judgments are made more intuitively could only be inferred indirectly from shorter reaction times and smaller ratings of
difficulty; and the group differences we found on the neural level did not directly support the idea that lawyers were indeed less emotional in their judgments. In general, our results call into question whether our stimulus material is apt to elicit differences in emotional processing and there could be various reasons for this. For example, we deliberately avoided to use dramatic cases like killing one’s own child in order to save the remaining family (Greene et al., 2001), we did not ask our subjects to imagine themselves as actors in these stories but to evaluate them from a “safe” distance, from an observer’s perspective, and adapted our case stories from more realistic incidents which had actually been reported in the media. If our cases thus failed to elicit emotions in the first place, then this would explain the absence of evidence suggesting a difference in emotional processing. In general, our results concerning normative judgment, that is, considering the similarities of moral and legal judgments taken together, suggest a great cognitive overlap of both kinds and emphasize the importance of social capacities such as the attribution of beliefs or intentions and mentalizing for normative cognition (Hauser, 2006; Huebner, Dwyer, and Hauser, 2009). In conclusion, the results of Study 1 suggest that there is at least some truth in the idealistic understanding of legal decisions and that there are more relevant findings regarding the way our brain processes normative judgment than can be understood when only investigating moral cognition.
Study 2 – Judging a Normative Judgment: An Exploration into Higher-Order Normative Cognition

With our second study we wanted to follow the trace of the engagement of theory-of-mind capacities during normative cognition a little further. The central idea was simple: Normative judgments can – at least in theory – be made on an indefinite number of levels. We can judge a situation directly, or we can judge a judgment of such a situation, and so on. In the first case, the object of the judgment is the situation itself, in the second it is the more abstract judgment of that situation. This idea also reflects the position of Jonathan Haidt who claimed that already engaging in normative discourse prompts normative judgments, even if people merely express their opinions on abstract and mediated ethical issues, that is, even if people discuss about moral judgments, not concrete moral situations (Haidt, 2001).

Our task design originally developed for Study 1 allowed us to investigate neural differences between direct and higher-order normative judgments, because in many cases the subjects did not have to judge whether a concrete situation was morally or legally right, but whether the judgment of such a situation – as it was made by an imaginary judge or a governmental institution – was morally or legally right. Thanks to the sufficient number of stimuli and the comparatively high number of experimental subjects, there were no methodological hindrances to examine these two kinds of normative judgments with a new statistical model.

A particular region of interest was the right temporo-parietal junction (rTPJ), whose importance for belief-attribution and theory-of-mind
capacities has been emphasized frequently in the neuroscientific literature (Aichhorn et al., 2009; Saxe and Kanwisher, 2003; Saxe 2006). If this aspect of social cognition is engaged differently in direct and higher-order normative judgment, then our comparison should yield different levels of activation within the rTPJ. This is indeed what we have found: The activation was relatively stronger during direct than during higher-order normative judgment. However, we must be aware that jumping from the consequent (that there is a difference in rTPJ activation) to the antecedent (that the engagement of theory-of-mind capacities for both kinds of judgment differs) would constitute the formal fallacy of affirming the consequent which is already expressed in the idea of the reverse inference addressed above. This emphasizes the importance of controlling for competing explanations of our results.

Fortunately, we could compare the different levels of normative judgment central to our experimental operationalization with regard to the categories of certainty, difficulty, reality, and emotional involvement, thanks to the subjects’ information we obtained in a post-experimental rating procedure. Particularly the former two categories are important to control for, as the essence of certain processes of social cognition is often difficult to grasp and unexpected differences in certainty or difficulty are likely to occur. Given the complex stimulus material employed in such experiments, it is often difficult, if not impossible, to change only the one aspect central to the respective investigation without influencing others. This difficulty is due to the intermixture of aspects of the social with different emotions, self-related
features, aspects of autobiographic memory, and so on, but even with lower-
level cognitive features like attention and control.

In our case, there were no differences in perceived certainty or difficulty, but small, yet significant differences in the perceived reality and emotional involvement between both categories. Additional correlation analyses were performed to see whether any of the behaviorally measured categories corresponded to the neural activation within the rTPJ but this was not the case. Since the rTPJ does not belong to those brain areas frequently associated with emotion, we defended the view that our results suggested the stronger engagement of theory-of-mind capacities during direct than during higher-order normative judgment. But we are well aware that there might be another feature related to the content of our stimulus material, one that we did not control for and that could account for the difference in brain activation. This demonstrates that future research is important to further investigate the role that theory-of-mind capacities play for normative judgment. However, our study provides some evidence that the level on which a normative judgment is made is of significance to those accounts of moral psychology emphasizing the attribution of beliefs and intentions (Hauser, 2006; Huebner, Dwyer, and Hauser, 2009) – a possibility that has been neglected by its proponents hitherto.
Study 3 – Moral Physiology, Its Limitations and Philosophical Implications

The summary of the previous two experiments already foreshadows one of the conclusions of Study 3, namely, that our knowledge about our moral physiology is still preliminary. Philosophers of science already noted how rare an experimentum crucis is in real science and it should thus not be surprising that the findings presented here only suggest the presence of some cognitive or emotional processes during normative decision making and inspire questions for future research. My review of the studies performed by other researchers, particularly those employing the moral dilemma task which was inspired by philosophical discussions, comes to a similar result. While the data are relevant to neuroscientific or psychological theories about brain functions or the presence of cognitive processes, it is less clear how they can be related to philosophical theories, for example, theories in ethics. One influential but, as I argued, failed example is Greene and colleagues’ contention that utilitarian judgments are more rational and less subject to intuitions than deontological judgments (Greene et al., 2004; Singer, 2005). This is particularly salient since the researchers did not operationalize their stimulus material properly in order to warrant any such conclusions (see below as well as Kahane and Shackel, 2008).

In a second step, I used the lesson previously learned from the example of moral physiology to address some general issues which are relevant to philosophy of mind, philosophy of science, and cognitive neuroscience. For example, Moll and colleagues’ conclusion that the studies of moral physiology show that a “remarkably consistent network of brain regions is
involved in moral cognition” (2005, p. 799) can actually be understood as a problem instead of a benefit of this research. If neuroscientists investigated moral phenomena in a number of different ways, that is, employing varying tasks (e.g. passive perception, reasoning, and judgement) and task modalities (e.g. visual text, visual images, auditory), and still found similar results as claimed by Moll and colleagues, then at least some important aspects of moral cognition are still neglected. This diagnosis lead to the analysis performed in section III.3 of Study 3, which addressed one question that has been frequently overlooked in cognitive neuroscience, namely, the granularity of cognitive as well as the underlying neural processes and in how far this has been reflected in current methods and tools of neuroimaging. It is an essential issue to clarify how a particular cognitive process can actually be individuated conceptionally and how this individuation can be operationalized in such a way that it can be related to neural processes – one of the essential tasks of contemporary cognitive neuroscience. In my opinion, we still need more conceptual as well as empirical work to really understand in which temporal, spatial, and physiological dimensions we have to ask the questions in our neural experiments in order to find answers we can also understand in cognitive or even philosophical terms.

More philosophical considerations exemplified by the example of moral physiology were addressed in a third step. Recurrent topoi which can be found abundantly in explanations within the field of cognitive neuroscience concern the distinction between emotional and cognitive processes. In this regard, moral physiology and psychology are no exceptions. The important
philosophical lesson to learn consists in the discovery that even if neuroscience provided results indicating that actual moral judgments are essentially subject to emotion but not to reason, or vice versa, there would still be the necessity to argue for the position that moral judgments should be made according to either of these views. Either empirical result could be just as well understood as motivation for a pedagogic mission to change the way people are making moral judgments – or how to cope with it if this should not be possible. These results also exemplify how useful, if not mandatory, it is to distinguish descriptive from normative issues. Finally, to the extent that debates in metaethics are necessarily debates about the meaning of concepts, it is not at all clear how neuroscientific investigations of moral phenomena could provide any evidence relevant to such discussions.

These considerations demonstrate, in my opinion, that there are genuinely philosophical tasks which cannot be taken over by research in neuroscience (see also Joyce, 2008). Even if many of my conclusions may be negative in character, they also emphasize positively the importance of interdisciplinary research by demonstrating what philosophers can learn from neuroscientists and, conversely, what neuroscientists can learn from philosophers. It is this insight which confirms the necessity to join the perspectives of both disciplines as was intended with the research comprising “Norms and the Brain”. 
Study 4 – The risk that neurogenetic approaches may inflate the psychiatric concept of disease and how to cope with it

The final exemplification of the combination of both perspectives within this dissertation consists in Study 4 and extends its scope into normative neuroethics and philosophy of psychiatry. It develops a thought experiment to demonstrate the risk that the psychiatric concept of disease could be inflated by future research within the fields of neuroimaging, neurogenetics, and the combination of both. Even though it is a thought experiment, it has a realistic background. One part of it is provided by the recent successes of “genomic imaging”, as the combination of neuroimaging and neurogenetics has been called (Caspi and Moffitt, 2006). The perspectives of both research methodologies taken in isolation suffer from intense limitations. While genetics captures information which remains constant throughout a subject’s lifetime, it is blind to the many phenotypical variations which are caused by environmental and epigenetic influences. Neuroimaging, by contrast, captures a present snapshot that is subject to such influences, but also highly variable due to many contingent factors influencing neural processes – many of which are irrelevant to the investigated issue and thus generate a higher degree of noise. There was hope that the combination of both perspectives, the highly constant, though somewhat outdated genetic, and the highly variable, though temporally more recent neural one might offer new insights into psychiatric disorders – and indeed this hope proved to be justified (e.g. Caspi and Moffitt, 2006; Meyer-Lindenberg and Weinberger, 2006).
The second respect in which the thought experiment has a realistic background cannot be found within science, but within incidents on the pharmaceutical market. One company, Cephalon, producer of the successful medication for narcolepsy and other sleeping disorders, recently proposed a new conceptual scheme of such disorders to the Foods & Drugs Administration (FDA) of the United States of America in order to increase the number of clinical conditions for which its medication is approved (Vastag, 2004). This included the controversial case of “subwakefulness syndrome”; it is controversial because it may classify a normal condition, namely, natural fatigue, as a disorder. The FDA did not accept Cephalon’s scheme. However, the thought experiment explores the possibility that such a scheme could include data from neuroimaging or even genomic imaging which was tailored by a clever selection of subjects in order to suggest that there is a biological cause – a so-called endophenotype – which corresponds to the conceptual scheme.

That such an approach is at all conceivable is subject to the nature of psychiatric disorders. They are inherently normative, often conceptually ambiguous and fuzzy, and even historically and transculturally variable. What is considered as mentally healthy and what as ill changes throughout time, between societies, and may even be controversial between experts within a particular society at one certain time. It is understandable if psychiatrists, under such conditions, look for better possibilities to define, diagnose, and distinguish mental disorders. Nevertheless, the thought experiment demonstrates that the decision about what is normal and what is ill can at best be informed or supported but never replaced by empirical
investigations alone. My conclusion emphasizes the importance of evaluating the validity of any such attempt, to apply one’s common sense and perhaps even the formalized means of the inference to the best conclusion, and that current research paradigms might carry the risk to look for solutions only on the neural or genetic level and thus neglect answers existing on the social level.

Summary and Outlook

The investigation into “Norms and the Brain” began with an issue that is central to our human self-understanding, namely, our capacity for normative judgment, and it also ended with such an issue, namely, the distinction between healthy and ill mental conditions. Both are related to what a human being is, and perhaps even to what a human being should be. The data analyses performed in Studies 1 and 2 provided some new answers on normative cognition but also raised many new questions. Future studies could, for example, employ cases which are more likely to elicit emotional responses during moral and legal judgment; could vary the legal content between the cases in order to better understand how expert knowledge interacts with these features; could keep the normative content constant when switching between direct and higher-order normative judgment, as was done between moral and legal decisions; could investigate different kinds of experts with normative knowledge; or could explicitly operationalize a certain kind of moral judgment to have an applicable standard for right and wrong. To somebody not familiar with empirical research the current state might appear frustrating – but even the thoughts
on improvements of future research constitute new insights and could hardly have been made without these experiments out of the armchair.

The philosophical analyses in Study 3 emphasized the importance of the scientific work that is necessary prior to any experiment, that is, at the stage of conceptualization, abstraction, and reduction in order to operationalize a theoretical question. The thoughts on the individuation and granularity of cognitive processes demonstrated the value of cognitive modeling – an exercise often neglected in neuroimaging but increasingly supported (and demanded) by computational neuroscience. One open question concerns the inferential strategies within neuroimaging research, whether and in how far the reverse inference problem can be coped with. Additional measures, such as psychological tests or conceptual analyses, are necessary to establish the presence of a cognitive process; but it might well be that a good degree of certainty requires so many precautions that neuroimaging will be degraded to a tool of validating – or rejecting – psychological predictions. However, these analyses also suggested that not all questions can be solved, perhaps some not even addressed, on the empirical level but need rigorous conceptual analyses; some might be essentially normative and subject to rational discussion and a good degree of common sense, as has been suggested in Study 4.

Scientific abstractions and reductions are necessary and the chosen operationalization to address an issue empirically may never be perfect. This does not generally diminish the value of scientific knowledge but it should remind the researchers of the epistemic status of their theories, particularly
at the stage of interpreting and explaining the findings. If one investigates only a tiny part of the mosaic, one has to be careful when making judgments about the whole piece. The presentation of neuroscientific results in the media, an issue that surpasses the scope of this investigation (but see Schleim, 2009), too often neglected these basic insights of modesty. It is arguable to what extent this falls into the responsibility of the researchers and to what extent into that of the journalists or editors. But even if opinions on this ethical issue might differ, it is important for the integrity of a so-called knowledge society to acknowledge, in the old Socratic tradition, what we do not (yet) know, to acknowledge the ignoramus. Such an exercise might even provide a remedy against the epistemic hubris that one’s current point of view is already at the end, or at least close to, the omega-point of ultimate knowledge. Concurrently, the justified fascination which is radiating from the pioneer work in neuroscience might provide a remedy against a pessimistic ignorabimus which separates empirically addressable issues from the domain of the knowable.3

References


3 The formats of the following studies have partially been modified in order to fit into this dissertation but the text corresponds to the final and published articles, where the texts are already published. That is, page numbers are continuous but numbering of footnotes, illustrations, and so an, as well as citation styles are individual for each study.


Study 1

From moral to legal judgment

The influence of normative context in lawyers and other academics

Stephan Schleim, Tade M. Spranger, Susanne Erk, Henrik Walter

Publication Note

At the time of submitting and defending my Ph.D. dissertation, this paper was under peer review. In the meantime, it was accepted for publication in Social Cognitive and Affective Neuroscience, published online on March 1, 2010 (doi:10.1093/scan/nsq010) and in press in 2011, Volume 6: 48-57.

As part of the revision process, the paper benefited from several of the reviewer's comments. These changes include four additional correlation analyses to further investigate the influence of legal expertise, an extended discussion of the relative lack of differences in brain processing between lawyers and other experts, and clarifications in the discussions of rule application as well as the function of the superior temporal gyrus.
Study 2

Judging a Normative Judgment

An Exploration into Higher-Order Normative Cognition

Stephan Schleim, Susanne Erk, Henrik Walter

Extended Abstract

Normative judgments about social customs as well as legal or moral issues are prevalent in everyday life. As research in moral psychology has shown, normative questions are relevant to us even if we are not directly involved, for example, when judging normative behavior within other cultures (Haidt et al., 1993). It has been argued in the so-called social intuitionist model of moral cognition that mere talking about normative attitudes already provokes normative decisions (Haidt, 2001). It is thus one outstanding feature of normative judgments that they cannot only be made about certain actions, whether imaginary or factual, but also about judgments of such actions, judgments of such judgments, and so on indefinitely, always expressing a subject’s stance on some normative issue.

In this explorative analysis of fMRI data recorded during a normative judgment task comprising moral and legal decisions (Schleim et al., 2010), we compared cases where subjects either directly made a judgment about a certain action or on a higher order about a judgment of such an action. In the latter case, the objects of the subjects’ decisions were judgments of third parties, such as judges or governmental institutions. Because the attribution of beliefs and intentions has been suggested as one primary cognitive requirement of normative judgments (Hauser, 2006; Huebner et al., 2009),
we wanted to investigate neural differences related to this capacity during
direct and higher-order normative judgment.

One of the brain areas consistently associated with the attribution of beliefs
and intentions, sometimes also referred to as “theory-of-mind” or
“mentalizing”, is the right temporo-parietal junction (rTPJ; Aichhorn et al.,
2009; Perner et al., 2006; Saxe, 2006; Saxe and Kanwisher, 2003). Since the
object of a higher-order normative judgment is itself a judgment as
contrasted to the object of a direct normative judgment which is a behavior
of human actors, we expected that the latter require more mentalizing
capacities than the former. In these cases, subjects had to reason about the
actors’ beliefs and intentions, while they were explicit in higher-order
judgments, where the actors clearly expressed their normative decision. We
wanted to test this hypothesis with the aid of our normative task design and
expected stronger activation of the rTPJ during direct than higher-order
normative judgment.

We re-analyzed the data of the 40 legal experts and other academics of the
previously published experiment on moral and legal decisions (Schleim et
al., 2010). Whereas the kind of normative judgment (moral or legal) was the
question of interest in the previous analysis, the level of normative judgment
(1st / direct or 2nd / higher-order) was at the focus of this exploratory
analysis. The direct judgments were about cases prompting the subjects to
evaluate a certain normative situation, whereas the higher-order judgements
were about cases prompting the subject to evaluate a certain decision,
typically of a judge or governmental institution, about a certain normative
situation. We identified 32 cases as prompting direct and higher-order normative judgements respectively and compared 16 of them in each group against each other. Further methods are analogue to the description in Schleim et al. (2010).

Analyses of behavioral data identified significant effects of level of judgment on reaction time as well as self-assessed measurements of reality and emotion (all \( p < 0.01 \)). Subjects required significantly more time during direct judgments, perceived themselves as emotionally more involved, but considered these cases as less realistic. Analyses of the fMRI data revealed significant differences of brain activation in the precuneus \( (3 / -44 / 42) \) and right temporo-parietal junction \( (51 / -51 / 13) \). Inspection of the mean beta values per cluster revealed a relatively stronger activation (less deactivation) in both areas for the direct than for the higher-order judgments. Correlation analyses between reaction time and self-assessment measures in the two clusters of brain activation did not reach significance.

According to the view so called “Rawlsian” views of moral cognition emphasizing the importance of action analyses for normative judgments (Hauser, 2006; Huebner et al., 2009), an eliciting situation triggers an analysis concerning who did what to whom for which reason or, in different terms, a causal-intentional analysis of that situation, in order to process a normative decision. However, when judging a normative judgment (i.e. when performing a higher-order normative judgment), the respective belief – that is, that the respective action is right or wrong – is explicitly expressed and thus does not have to be inferred from the situation. It was thus
reasonable to expect differences in brain activation when comparing direct normative judgments, which have already been investigated previously, with higher-order normative judgments, which have been neglected hitherto.

Although areas within the prefrontal cortex have been associated with mentalizing and theory-of-mind previously (Amodio and Frith, 2006; Van Overwalle, 2009), it has been argued that the more specific capacity of attributing beliefs and intentions to others is subject to the TPJ, particularly the rTPJ (Aichhorn et al., 2009; Saxe, 2006). According to this interpretation, our data suggest that this social capacity is indeed more important for direct than for higher-order normative judgments. The exploratory analysis based on our conceptual distinction between different kinds of levels of normative judgments was thus able to identify a difference in brain activation in precisely this area which is believed to be central for the essentially social processes of attributing beliefs and intentions.

This finding is relevant to the current debate in moral psychology discussing the essential role of causal-intentional analyses in moral cognition, such as proposed by the “Rawlsian” model (Hauser, 2006; Huebner et al., 2009). According to our findings and our interpretation, normative judgments can be of different kinds which require different amounts of theory-of-mind capacities. Thus, a strictly “Rawlsian” approach toward moral cognition carries the risk of neglecting other cognitive processes which may also be relevant to normative cognition.


Study 3

Moral Physiology

*Its Limitations and Philosophical Implications*

Stephan Schleim

Publication Note

This paper is published in 2008 in *Jahrbuch für Wissenschaft & Ethik* 13: 51-80.
Study 4

The risk that neurogenetic approaches

may inflate the psychiatric concept of disease and how
to cope with it

Stephan Schleim

Publication Note

This paper is published in 2009 in Poiesis & Praxis 6: 79-91.