

# Moral Neuroscience and Moral Philosophy: Interactions for Ecological Validity

Koji TACHIBANA\*

## Abstract

Neuroscientific claims have a significant impact on traditional philosophy. This essay, focusing on the field of moral neuroscience, discusses how and why philosophy can contribute to neuroscientific progress. First, viewing the interactions between moral neuroscience and moral philosophy, it becomes clear that moral philosophy can and does contribute to moral neuroscience in two ways: as *explanandum* and as *explanans*. Next, it is shown that moral philosophy is well suited to contribute to moral neuroscience in both of these two ways in the context of the problem of ecological validity. Philosophy can play the role of an *agent for ecological validity*, since traditional philosophy shapes and reflects part of our social reality. Finally, based on these arguments, I tentatively sketch how a Kantian account of moral incentive can play this role.

## 1. Introduction

There are many approaches to think about the relationship between neuroscience and society. On the one hand, for example, we can think about the Ethical, Legal, and Social Issues/Implications (ELSI) of neuroscience. This is called the “ethics of neuroscience” or “neuroethics” in its narrow sense (Illes (ed.) 2006). On the other hand, we can also think about how we govern neuroscience itself, i.e., “neurogovernance” (Tachibana 2009a, 2009b). Philosophy, as well as bioethics, law, politics, and STS (Science, Technology, and Society), have all contributed to these fields.

---

\* E-mail: tachibanakoji.philosophy@gmail.com

There is, however, another approach philosophy can take to that relationship, which is to contribute to the progress of neuroscience itself. This essay discusses how and why philosophy can contribute to neuroscientific progress. For the sake of clarity, I will focus on the field of moral neuroscience. First, I survey moral neuroscience, its methods and claims (section 2). I then discuss two ways in which moral philosophy contributes to moral neuroscience (section 3 and 4). Third, I discuss why moral philosophy can contribute to moral neuroscience in these two ways, focusing on them from the viewpoint of ecological validity. Specifically, I propose that philosophy can play the role of an agent for ecological validity since traditional philosophy shapes and reflects part of our social reality (section 5). Finally, based on these arguments, I tentatively sketch how a Kantian account of moral incentive can play this role (section 6).

## **2. Moral Neuroscience**

What does moral neuroscience tell us about the nature of morality? Neuroscientists often start with the stories of Phineas Gage and Patient Elliot who displayed emotional, social, and moral behavioral problems despite possessing normal or above par intellectual faculties (A.R. Damasio 1994). What is interesting for neuroscientists is that these symptoms appeared following damage to the ventromedial prefrontal cortex (VMPFC) or the orbitofrontal cortex (OFC) (H. Damasio *et al.* 1994). It later became clear that other VMPFC/OFC damaged patients showed similar behavioral traits (A.R. Damasio 1994; Amari and Kato (eds.) 2008, 10ff.). This suggests two things; that emotion may have an important connection to morality, and that the VMPFC/OFC may play a crucial role in that connection.

Recent research shows that there are many neural correlates of morality. In the cerebral cortex, the dorsolateral PFC (DLPFC), the superior temporal sulcus region (STS), the anterior temporal lobe (ATL), the anterior PFC, the OFC, and the VMPFC; in the subcortex, the amygdala, the hypothalamus, the basal forebrain, and the septal area (Moll *et al.* 2008a). The anterior cingulate cortex (ACC) and the right temporoparietal junction (RTPJ) also play some role (Greene *et al.* 2004; Young *et al.* 2007).<sup>1</sup> Although there are many differences in the details of the models (cf. A.R. Damasio 2003; Greene 2007; Moll and de Oliveira-Souza 2007b), what they have in common is the idea that two systems are necessary for moral behavior; one is an emotional/intuitive system mainly

based on the VMPFC, the OFC, and the cerebral limbic system, and the other is a rational/cognitive system mainly based on the DLPFC (Greene 2007; Koenigs *et al.* 2007; Moll and de Oliveira-Souza 2007a; Greene *et al.* 2008; Moll *et al.* 2008a; cf. Zelazo *et al.* 2005). Thus it is a claim of moral neuroscience that human beings need emotion as well as reason to be moral.

For the purposes of this essay, I would like to note two points about this neuroscientific claim. First, it is in contradiction with the traditional understanding of morality, especially the Western tradition, which tends to regard morality as governed by the faculty of human rationality (as many philosophers and psychologists have maintained, including Plato, Aristotle, Kant, John Stuart Mill, Piaget, and Kohlberg),<sup>2</sup> and which traditional moral theories have historically shaped and reflected. I will return to this point later, although I think the fact that the neuroscientific claims have an enormous impact on society partly supports this point.<sup>3</sup>

The other point I would like to make concerns the methodology of neuroscience. Neuroscience must base its arguments on experimental evidence because it is an experimental science. Experiments must satisfy various conditions: they are based upon such concepts as “operational definition”, “manipulation”, and “control”; they are modulated by statistical significance-detection and controlled variables; and they are performed with the use of experimental devices such as functional MRI and PET. These conditions are necessary, though not sufficient, for well-designed experiments. For a well-designed experiment is an experiment which is, in addition, *ecologically valid*. An ecologically valid experiment is an experiment whose tasks or situations reflect the real world, with all of its complications, in which the subject of the experiment is normally embedded. Therefore the more ecologically valid an experimental design is, the better the experiment is.<sup>4</sup>

If traditional moral theories shape and reflect part of our culture, society, and tradition, and if moral neuroscientists are to perform ecologically valid experiments, then we can expect that moral philosophy can contribute to moral neuroscience. In the following sections, we will see two examples that demonstrate this contribution.

### 3. Moral Philosophy as *Explanandum*

If an experiment can examine the nature of morality *via* traditional moral

theories, it might be considered a well-designed experiment since these theories have shaped and reflected part of our culture, society, and tradition. Some neuroscientists seem to think along these lines when they refer to moral theories in their articles: “[a]ccording to classical moral theories, moral behavior is a perfectly rational type of affair” (Ciaramelli *et al.* 2007, 84); “[t]hrough the centuries, philosophical theories have adopted a deductive logico-verbal approach to morality that aims to identify universal principles that should guide human conduct” (Moll *et al.* 2005, 799); “[t]he long-standing rationalist tradition in moral psychology emphasizes the role of reason in moral judgment” (Greene *et al.* 2001, 2105). Here we can see one way in which moral philosophy can contribute to moral neuroscience. In this section, I will examine this further through two experiments; one by Greene *et al.* (2001, 2004) and the other by Koenigs *et al.* (2007).

To research the neural correlates of morality, Joshua Greene and others use a so-called “trolley problem” first described by the philosopher Philippa Foot (1967). They modified Foot’s original problem so as to highlight a conflict between deontological and utilitarian judgment.<sup>5</sup> The modified trolley problem is constituted by two dilemmas; one is known as an “impersonal” moral dilemma and the other is known as a “personal” moral dilemma. The impersonal dilemma is a task asking whether it is better to push a button to turn a runaway trolley away from five people and toward a single person. In this dilemma, utilitarian principles morally require us to push the button, killing one person but saving five. The personal dilemma is a task asking whether it is better to push a fat person off a bridge (killing him) in order to stop the runaway trolley from hitting and killing five people. In this dilemma, deontological principles morally require us not to push him. The result shows that when people make a deontological judgment, the emotion-related regions activate significantly more than when they make a utilitarian judgment. This suggests that deontological judgment is emotional rather than rational.

Michael Koenigs and others (2007) followed up on Greene’s approach. To investigate whether emotions mediated by the VMPFC have a crucial influence on moral judgments, they studied the moral judgment of the VMPFC damaged patients. Their experiment shows that these patients exhibit an abnormally high rate of utilitarian judgments on the high-conflicted personal dilemmas, which are emotionally salient and concern with aggregate welfare. This suggests that

a sort of emotion mediated by the VMPFC is needed to perform deontological moral judgments (see also Ciaramelli *et al.* 2007).

These experiments indicate that deontological judgments are emotion-laden and that the VMPFC is a neural basis for these judgments. Some neuroscientists have concluded that this evidence also shows that deontological moral theories are not ecologically valid: Greene (2008, 36-37) writes that “if these empirical claims are true, they may have normative implications, casting doubt on deontology as a school of normative moral thought” and that therefore he “believe[s] it is possible that philosophers do not necessarily know what consequentialism and deontology really are”. They take such a *negative or skeptical* attitude toward moral philosophy and use moral philosophy as *explanandum* which will be eventually be explained away.

Some philosophers have taken this idea on board. For example, responding to Greene, Mark Timmons (2008, 102), a Kantian philosopher, writes that “[d]eontology is committed to the idea that moral judgments are beliefs or are more cognitive than the evidence shows us; in short, deontology is committed to moral rationalism. However, in the light of empirical evidence about people’s intuitive moral judgments, a nonrationalist, sentimentalist account of them is more plausible than rationalist accounts. Thus, deontology is mistaken”.<sup>6</sup> Philosophers identified with the “experimental philosophy” movement have also attempted to show that philosophical theories do not necessarily reflect lay intentions (Appiah 2008; Knobe and Nichols (eds.) 2008).

These attitudes suggest in contrast that moral theories are thought to have shaped and reflected part of our society and tradition for a long time even if the thought might be a delusion. For if we do not think so from the beginning, their experiments and attempts will lose face or their selling point. Therefore, here we can see one kind of contribution that moral philosophy can make. That is, moral philosophy can contribute to moral neuroscience as *explanandum*, since neuroscientists are able to examine the nature of morality *via* traditional moral theories.

#### **4. Moral Philosophy as *Explanans***

Although many philosophers oppose the negative attitude described in the previous section for various reasons, I suspect that the vast majority of moral neuroscientific experiments have been performed in the absence of a thorough

understanding of traditional moral theories. I expect, in other words, that current experiments may be able to be made more ecologically valid if moral theories are read more carefully. Here we can see another way in which moral philosophy can contribute to moral neuroscience. I will sketch out this way by surveying a series of experiments and arguments performed by Jorge Moll and others (2002a, 2002b, 2005, 2008a, 2008b) who have successfully revised their experiments by applying a traditional moral theory, *virtue ethics*.

In an attempt to find the neural correlates of moral emotion, which is distinct from basic emotions, Moll *et al.* performed two experiments using various social statements (2002a) and pictures (2002b). They categorized the statements/pictures into four/six types distinguishing, moral ones from non-moral ones. Thus the experimental designs distinguished between moral and non-moral social norms. The results showed that the anterior aspect of the MOFC, the MPFC, and the STS are significantly active when people are reading/viewing morally un-/pleasant social statements/pictures, compared with non-moral ones. Based on these experiments, they concluded that these regions relate to the moral emotions.

Although stimuli pertaining to their categorizations were selected based on behavioral studies in independent groups of subjects with instructions, as is the case with most neuroscientific experiments, we can question what kind of distinction between moral situation and social one the experimental designs assume. For example, in 2002a, the statement “he never uses the seat belt” is classified as a non-moral and relatively unemotional social situation. Although it clearly involves violating the social norm of acting within the law, it is unclear what kind of moral aspect is lacking in this socially illegal act.

We can see this problem more clearly in their 2002b where the International Affective Picture System (IAPS) was used. While the specific pictures used in the experiment cannot be identified, there are many highly social-dependent pictures within IAPS. For example, there is a picture of a large number of dollar bills (No. 8501), which is rated as highly pleasant and highly emotionally arousing. There is also a picture of the Ku Klux Klan burning a cross (No. 9810), which is rated as highly unpleasant and highly emotionally arousing.<sup>7</sup> These pictures have a similar rating to some emotionally arousing biologically pleasant/unpleasant pictures (e.g., a crushed bloody human), which suggests that the same IAPS pictures may well have different affective ratings in subjects

across different cultures. For example, although Japanese people will find some biologically disgusting pictures equally unpleasant, they will likely not have the same level of emotional response as American people when seeing a picture of dollar bills or of the KKK burning a cross. Moreover, Japanese people would likely even categorize the picture of the KKK burning a cross as *non-moral* social situation, while Americans would categorize it as *moral* one.

There is some evidence to support this suggestion. Rafaela Ribeiro and others (2005) compare Brazilian and American norms for IAPS and conclude that although the pleasure and dominance dimensions of IAPS pictures are appropriate for use in Brazil, “[t]he differences found in the subjective arousal levels induced by some positive photographs must be taken into account in studies that will employ the norms presented here and may lead to discussions about the meaning of the term ‘arousal’ as employed by Lang *et al.* [(2005)], which should be more clearly defined” (pp. 214-215). In addition, Joan Chiao and others (2008) claim that even for fear, one of the so-called basic emotions, the bilateral amygdala response to fearful faces is partly modulated by culture.

In general, it is plausible that the response elicited by some statements/pictures depends on the local context in which subjects have grown up, and therefore that Moll *et al.*’s classification may not hold universally but rather be society-dependent. Although this point does not directly speak against the scientific value of the experiments, it does naturally suggest worries about whether they assume universal morality (the Kantian view) or society-dependent morality (the Aristotelian view), and whether the experimental designs are appropriate to investigate what we mean by “morality”.

It is important for our purpose to note that Moll *et al.* have revised their experimental strategy after interaction with philosophers. Three years after the experiments described above, referring to a suggestion of philosopher, William Casebeer (2003), they write that “[e]cological validity is especially relevant for moral cognition studies, because moral cognition depends strongly on situational and cultural context” (Moll *et al.* 2005, 803). Three years later still, referring to another philosopher, Alasdair MacIntyre, a contemporary advocate of virtue ethics, they decide upon a new operational definition of morality: “[m]oral’ (derived from the Latin *moralis*) and ‘ethics’ (from the Greek *êthikos*) originally referred to the consensus of manners and customs (MacIntyre 1985). Based on this rather broad notion, we have operationally adopted the definition

of morality as the sets of customs and values that are embraced by a cultural group to guide social conduct” (Moll *et al.* 2008a, 161; 2008b, 2-3; see also 2005, 799: italics in original). This definition clarifies the relationship between morality and social norms, which Moll *et al.* later found to be relatively unclear in their original definitions.

This example shows us another way in which moral philosophy can contribute to moral neuroscience. The long history of moral philosophy allows it to inform the developing theories of moral neuroscience from multiple perspectives. These perspectives help neuroscientists to revise their experimental designs so as to reflect social reality, making possible new and radical avenues of research. If there is something to be said for this argument, then a skeptical attitude which regards moral philosophical theories only as *explananda* may be too hasty. In contrast, it may be possible to use them as *explanantia* to bring the multifacetedness of subjects to the attention of neuroscience and increase the ecological validity of the experiments. This recommends us to take a *positive or favorable* attitude toward moral philosophy.

## 5. Interactions for Ecological Validity

In the previous sections, we saw two kinds of contribution that moral philosophy can make to moral neuroscience. First, as *explanandum*: it is useful to examine the nature of morality *via* moral theories, since these theories are traditionally understood to be ecologically valid for our society. Second, as *explanans*: the multifaceted perspectives of moral philosophy give new perspectives on current experiments. Though the first contribution is based on the negative attitude toward moral philosophy and the second is on the positive one, they are both based on the fact that moral theories have shaped and reflected part of our culture, society, and tradition.

Moreover, they are compatible: on the one side, even the neuroscientists who take the positive attitude may find that some types of moral theories have poor ecological validity; on the other side, even those who take the negative attitude may have an opportunity to revise their experiments in the light of philosophical criticism. For example, Koenigs *et al.* (2007) has received criticism from philosophers that dilemmas such as the trolley problem do not demonstrate dilemmas between deontological and utilitarian judgments (Kahane and Shackel 2008). To demonstrate this, they sent questionnaires to five other



philosophers asking what kind of horns the dilemmas really have. The five philosophers classified the dilemmas into five horns; “impartial consequence (C)”, “duty not to harm (D)”, “egoistic self interest (SI)”, “consequence-involving-egoistic-self-interest (C/SI)”, “mixed consequence and desert (C/D)”. Based on this classification, they claim that if Koenigs *et al.* are right, VMPFC damaged patients endorse utilitarian judgments only in high-conflicted personal dilemmas (C vs. D). However, for example, some of dilemmas which Koenigs *et al.* categorize as low-conflicted are also C vs. D. Therefore, they argue, Koenigs’ experiment is not appropriate as a test for utilitarian judgments: “[i]n conclusion, to establish that a response pattern manifests a tendency to utilitarian moral judgement, the stimuli used need to be classified in terms of content and not by purely behavioural or emotional criteria as was done here and in other studies such as those of Greene *et al.* [(2001, 2004; Ciaramelli *et al.* 2007)].”

Neuroscientists reply to these philosophical criticisms as follows. Greene (2008, 38–39) writes that “I will put aside their conventional philosophical definitions and focus instead on their relevant functional roles.... This is a ‘characteristically consequentialist’ judgment because it is easily justified in terms of the most basic consequentialist principles” (see also Koenig *et al.* 2008).

Although this claim shows that Greene has adopted the negative attitude toward moral philosophy as we have already seen in section 3, it is important to note that his attitude is connected to the positive one, which enables neuroscientists to revise their current experiments. For example, to reply Kahane and Shackel’s criticism, it will be sufficient to contrive an experimental design which divides some philosophical and traditional terms such as “deontological” and “utilitarian” into further distinctions. Greene *et al.* (2009) have performed a new behavioral experiment along these lines and show that deontological judgment in personal moral dilemmas is constructed of both “personal force” and “intention”. Moreover, in another experiment which studied the pattern of neural activity associated with honest and dishonest moral decisions, Greene and Paxton (2009, 12510) acknowledge the limitations of this current experiment and the possibility for the further experiments, writing that “it is not known whether our task is an ecologically valid model for real-world lying”.

These philosophy-neuroscience interactions show that although it will and should depend on each neuroscientist the extent to which they incorporate philosophical perspectives, ecological validity cannot be side-stepped and external criticism, such as philosophical criticism, is useful to revise the current experiments into ecologically more valid experiments. In this way, negative and positive attitudes are compatible.

Tim van Gelder (1998), characterizing philosophical method as argument, conceptual clarification, and historical perspective, lists seven roles philosophers can play in cognitive science. Although I think philosophers can also play all of the roles within neuroscience, here I want to emphasize another role based on the discussion so far, namely the philosopher as *an agent for ecological validity*.<sup>8</sup> If moral theories, as a historical fact, have shaped and reflected our culture, society, and tradition, they are in this sense ecologically valid and therefore, they can contribute to moral neurosciences in various ways. In detail, the role of moral philosophy as *explanandum* is to *be accused* of our misunderstanding what is ecologically valid; while that of moral philosophy as *explanans* is to *accuse* neuroscientists of performing experiments with poor ecological validity, failing to reflect the multiple aspects or true form of our morality. Since both of these roles concern ecological validity, moral philosophy can play the role of an agent for ecological validity.

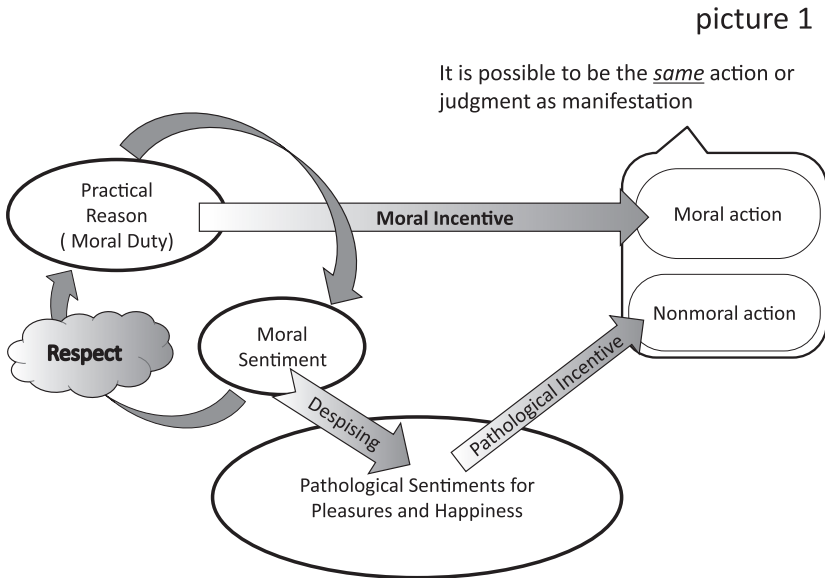
In the next, I attempt, albeit sketchily, to make concrete what the role entails.

## 6. A Kantian Account of Moral Incentive

Here I tentatively sketch how a Kantian account of moral incentive can be used as a tool for increasing ecological validity. It is usually said that Kantian deontology sets great value on human reason as a duty. Neuroscientists have criticized deontology for this reason. But Kant also discusses the moral incentive to act in accord with one's duties in the notorious third chapter of the *Critique of Practical Reason* (Kant 1788/1990), entitled "The Incentive of Pure Practical Reason," or we may say "A Theory of Moral Sensibility" (Beck 1960, 219; Allison 1990, 121).

In this chapter, Kant argues that what makes an action morally good is not the action itself but its incentive. He divides human incentives into two types: one is pathological, based on which we usually act; the other is moral, which makes our action morally good. He calls the latter "respect for moral law":

“Achtung fürs moralische Gesetz ist also die einzige und zugleich unbezweifelte moralische Triebfeder” (Kant 1778/1990, 139, 147). He also points to two aspects of moral incentive: to respect the moral law, and to despise pathological incentives (Verachtung). I schematize this Kantian picture of moral incentive below (picture 1):



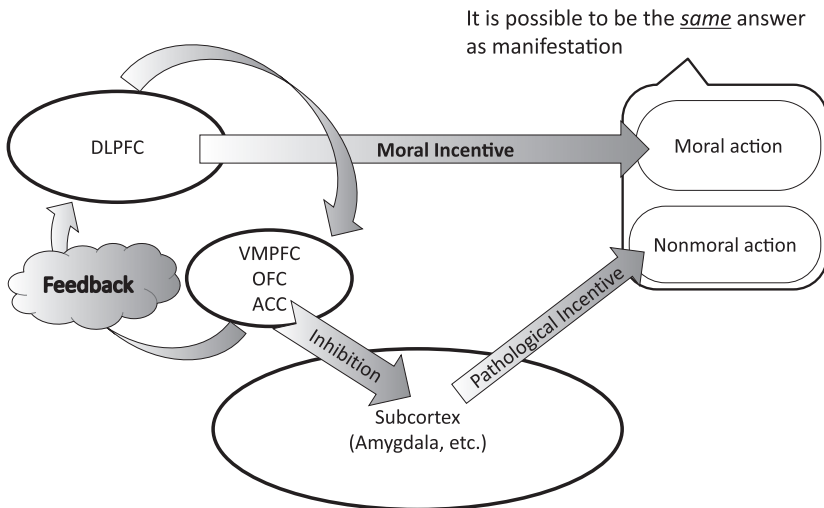
This picture seems to reflect our traditional understanding of morality. We do not regard a person who acts rightly as morally good if the person acts from a bad incentive.<sup>9</sup> Putting it in the context of neuroscientific experiments, this distinction indicates that fMRI studies on the responses to moral questionnaires may not provide the neural correlates of morality if the answers (e.g., “No, a fat person should not be pushed off a bridge to stop a runaway trolley from hitting five people”) are based on a morally bad incentive. Therefore, an experimental design which reflects this distinction will be ecologically more valid than designs which do not reflect it. This Kantian distinction encourages us to take a new approach to morality, one which does not rely solely on subjects’ “yes” or “no” response.

However, there is little point in asking subjects about the incentive or reason

for their answers, since it is known that subjects can make *post hoc* justifications (Haidt 2001), sometimes fail to provide justification at all for moral judgments (Hauser *et al.* 2007), and be easily biased by emotional context and social group (Valdesolo and DeSteno 2006, 2007). Neuroscientists, therefore, have employed two alternative methods for discovering this. One is to fractionate tasks (Greene *et al.* 2009). The other is to classify moral emotions and investigate them separately, e.g., honesty (Greene and Paxton 2009), envy/schadenfreude (Takahashi *et al.* 2009), and so on (Rozin *et al.* 1999; Haidt 2003; Moll *et al.* 2008b).

Here I propose what might be called a third method. As Kant shows, the same action or judgment could be either morally good or morally bad according to its incentive/reason. If we can detect goodness/badness of subjects' moral incentive *via* neural network of brain activation, we may not have to depend on their verbal reports. As we saw in section 2, there are many neural correlates of morality, with the VMPFC and the OFC playing an important role (see also Kringsbach and Rolls 2004, 354-355). Based on this, I propose the following model (picture 2):

picture 2



The point of this model is to try to find the neural basis of morality not by the single-detection of a simple correlation between subjects' answer and neural activity, but rather by multiple-detection of the *weighting* of activation patterns across different brain regions. This model may be effective for distinguishing between two subjects, one morally virtuous and the other not, who give the same right answer to a question. If there are different patterns of activation across regions (roughly speaking, e.g., three regions of the DLPFC, the VMPFC/OFC/ACC, and the subcortex) between the two subjects, we might speculate that the first answer is motivated by a morally good incentive, and the other by a morally bad incentive.

Although this model must be expanded in order for neuroscientists to make use of it to design new experiments (cf. Cunningham and Zelazo 2007), it depends on transdisciplinary discussion between philosophers and neuroscientists whether this model will bear fruit.<sup>10</sup>

## 7. Conclusion

I have discussed two kinds of contribution that philosophy can make to neuroscience, and proposed that philosophy can play the role of an agent for ecological validity. Although I have focused on moral philosophy in this paper, and psychology often plays such a role, various research areas within the humanities can also be expected to play a similar role, since the humanities in general have shaped and reflected our culture, society, and tradition.

Certainly, it will not be easy to perform an ecologically valid experiment for various reasons (Casebeer and Churchland 2003). There is still, I think, substantial philosophical work to be done in this new and engaging field. The value of a philosopher's achievement here will not be evaluated only by his or her conceptual analyses but also by experimental results.

## Acknowledgments

This essay is supported by a Grant-in-Aid for Scientific Research provided by the Japanese Ministry of Education, Culture, Sports, Science and Technology. I would like to thank Dr. Guy Kahane (University of Oxford) and Dr. Jorge Moll (Neuroscience Center, LABS-D'Or Hospital Network) for generously providing me with further details of their experiments, Prof. Hideki Ohira (Nagoya University) for his informative comments, and Dr. John O'Dea (University of Tokyo) for his help in editing my English.

## Notes

1. These regions are said to relate not only to morality but also to many other functions such as working memory, attention, executive function, and goal-directed behavior (Doya 2008; Osaka (ed.) 2008). With respect to goal-directed behavior, according to one model the OFC responds to rewards as a goal, the DLPFC formulates a plan to get the reward, and the MPFC initiates the appropriate action (Amari and Kato (eds.) 2008, 23). Antonio R. Damasio (1994, 2003) proposed the well-known “somatic-marker hypothesis” model, according to which the VMPFC combines external stimuli and emotion, and causes somatic responses.
2. Although there are many differences between the Western and Eastern intellectual traditions (see section 4), there are also many influences, as Endo (2006, 304) points out: “We have focused on the role of reason for a long time. For we have been influenced by the tradition of western philosophy in which philosophers such as Plato and Kant assume the confrontation between reason and passion”.
3. Even in Japan, the neuroscientific claim has received full coverage in some nationwide newspapers (April. 20, 2007 in *The Mainichi Dairy News*; April. 11, 2008 in *The Asahi Shimbun*).
4. The argument of this essay does not, however, depend on the difference between ecological validity and external validity.
5. This modification itself has been done in various ways by philosophers and psychologists.
6. Although Timmons himself is “generally sympathetic to sentimentalism”, he thinks “deontology *per se* is not threatened by the empirical work cited by Greene; there are versions of deontology that can avoid Greene’s arguments” (Timmons 2008, 102-103).
7. In the version of 2005 (Lang *et al.* 2005), No. 8501 is rated as Valence Mean 8.14, Arousal Mean 6.86, and No.9810 as VM 2.25, AM 6.74. (Moll *et al.* 2002b uses the version of 1995.)
8. Gelder (1998, 125-126) explicitly writes that the list is not intended as definitive, and notes that “philosophers contribute to cognitive science in many and diverse ways, and any reader familiar with cognitive science can probably think of interesting additions or alternatives”.
9. There is no space here to discuss the relationship between “morally right” and “morally good” (cf. Stratton-Lake 2006).
10. There are two things worth noting here: first, this model seems to be compatible in its approach with some current empirical studies. For example, Greene *et al.* (2009) show that the bilateral-DLPFC and the ACC activate when a dishonest person give an honest answer, but do not activate when an honest person gives an

honest answer. Second, the German word “Achtung” usually means “attention” as well as “respect”, although there is a controversy on its philosophical interpretation (cf. Paton 1948, 63–64; Allison 1990, 125–128). If this concept, i.e., respect as moral incentive, has a moral cognitive aspect as well as a moral emotional aspect in a Kantian account of moral incentive, philosophers may be able to shed light on the interaction between moral cognition and moral emotion, which is currently a topic of lively debate in moral neuroscience.

## References

- Allison, H.E. 1990. *Kant's Theory of Freedom*. Cambridge University Press.
- Amari, S. and T. Kato. (eds.) 2008. *Neuroscience of the Mind: Implications from Biological Psychiatry*. University of Tokyo Press. (Amari is the supervising editor: Japanese.)
- Appiah, K.A. 2008. *Experiments in Ethics*. Harvard University Press.
- Beck, L.W. 1960. *A Commentary on Kant's Critique of Practical Reason*. University of Chicago Press.
- Casebeer, W.D. 2003. Moral Cognition and Its Neural Constituents. *Nature Reviews Neuroscience*, 4:841–846.
- Casebeer, W.D. and P.S. Churchland. 2003. The Neural Mechanism of Moral Cognition: A Multi-Aspect Approach to Moral Judgment and Decision-Making. *Biology and Philosophy*, 18:169–194.
- Chiao, J.Y., T. Iidaka, H.L.Gordon, J. Nogawa, M. Bar, E. Aminoff, N. Sadato, and N. Ambady. 2008. Cultural Specificity in Amygdala Response to Fear Faces. *Journal of Cognitive Neuroscience*, 20(12):2167–2174.
- Ciaramelli, E., M. Muccioli, E. Làdavas, and G. di Pellegrino. 2007. Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex. *Social Cognitive and Affective Neuroscience*, 2:84–92.
- Cunningham, W.A. and P.D. Zelazo. 2007. Attitudes and Evaluations: A Social Cognitive Neuroscience Perspective. *Trends in Cognitive Science*, 11(3):97–104.
- Damasio, A.R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. G.P. Putnam.
- . 2003. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Harcourt.
- Damasio, H., T. Grabowski, R. Frank, A.M. Galaburda, and A.R. Damasio. 1994. The Return of Phineas Gage: Clues about the Brain from the Skull of a Famous Patient. *Science*, 264:1102–1105.
- Doya, K. 2008. Modulators of Decision Making. *Nature Neuroscience*, 11(4):410–416.
- Endo, T. 2006. Emotion. In: H. Kaiho and T. Kusumi. (eds.) 2006. *The Comprehensive*

- Dictionary of Psychology*. Asakura-shoten, pp. 304-334. (Japanese.)
- Foot, P. 1967. The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, 5:5-15.
- van Gelder, T. 1998. The Roles of Philosophy in Cognitive Science. *Philosophical Psychology*, 11(2):117-136.
- Greene, J.D. 2007. Why Are VMPFC Patients More Utilitarian? A Dual-Process Theory of Moral Judgment Explains. *Trends in Cognitive Science*, 11(8):322-323.
- . 2008. The Secret Joke of Kant's Soul. In: W. Sinnott-Armstrong. (ed.) 2008. *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. MIT Press, pp. 35-79.
- Greene, J.D., F.A. Cushman, L.E. Stewart, K. Lowenberg, L.E. Nystrom, and J.D. Cohen. 2009. Pushing Moral Buttons: The Interaction between Personal Force and Intention in Moral Judgment. *Cognition*, 111:364-371.
- Greene, J.D., S.A., Morelli, K. Lowenberg, L.E. Nystrom, and J.D. Cohen. 2008. Cognitive Load Selectively Interferes with Utilitarian Moral Judgment. *Cognition*, 107:1144-1154.
- Greene, J.D., L.E. Nystrom, A.D. Engell, J.M. Darley, and J.D. Cohen. 2004. The Neural Bases of Cognitive Conflict and Control in Moral Judgment. *Neuron*, 44:389-400.
- Greene, J.D. and J.M. Paxton. 2009. Patterns of Neural Activity Associated with Honest and Dishonest Moral Decisions. *Proceedings of the National Academy of Sciences*, 106(30):12506-12511.
- Greene, J.D., R.B. Sommerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen. 2001. An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, 293:2105-2108.
- Haidt, J. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108(4):814-834.
- . 2003. The Moral Emotion. In: R.J. Davidson, K.R. Scherer, and H.H. Goldsmith. (eds.) 2003. *Handbook of Affective Sciences*. Oxford University Press, pp. 852-870.
- Hauser, M., F. Cushman, L. Young, R. Jin, and J. Mikhail. 2007. A Dissociation between Moral Judgments and Justifications. *Mind and Language*, 22(1):1-21.
- Illes, J. (ed.) 2006. *Neuroethics: Defining the Issues in Theory, Practice, and Policy*. Oxford University Press.
- Kant, I. 1778/1990. *Kritik der praktischen Vernunft*. Meiner.
- Kahane, G. and N. Shackel. 2008. Do Abnormal Responses Show Utilitarian Bias? *Nature*, 452:E5.
- Knobe, J. and S. Nichols. (eds.) 2008. *Experimental Philosophy*. Oxford University Press.



- Koenigs, M., L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, and A.R. Damasio. 2007. Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments. *Nature*, 446:908-911.
- . 2008. Koenigs *et al.* Reply. *Nature*, 452:E5-E6.
- Kringelbach, M.L. and E.T. Rolls. 2004. The Functional Neuroanatomy of the Human Orbitofrontal Cortex: Evidence from Neuroimaging and Neuropsychology. *Progress in Neurobiology*, 72:341-372.
- Lang, P.J., M.M. Bradley, and B.N. Cuthbert. 2005. *International Affective Picture System (IAPS): Digitized Photographs, Instruction Manual and Affective Ratings. Technical Report A-6*. University of Florida, Gainesville, FL.
- MacIntyre, A.C. 1985. *After Virtue: A Study in Moral Theory*. 2nd ed., Duckworth.
- Moll, J. and R. de Oliveira-Souza. 2007a. Moral Judgments, Emotions and the Utilitarian Brain. *Trends in Cognitive Science*, 11(8):319-321.
- . 2007b. Response to Greene: Moral Sentiments and Reason: Friends or Foes? *Trends in Cognitive Science*, 11(8):323-324.
- Moll, J., R. de Oliveira-Souza, I.E. Bramati, and J. Grafman. 2002(a). Functional Networks in Emotional Moral and Nonmoral Social Judgments. *Neuroimage*, 16:696-703.
- Moll, J., R. de Oliveira-Souza, P.J. Eslinger, I.E. Bramati, J. Mourão-Miranda, P.A. Andreiuolo, and L. Pessoa. 2002(b). The Neural Correlates of Moral Sensitivity: A Functional Magnetic Resonance Imaging Investigation of Basic and Moral Emotions. *The Journal of Neuroscience*, 22(7):2730-2736.
- Moll, J., R. de Oliveira-Souza, and R. Zahn. 2008(a). The Neural Basis of Moral Cognition. *Annals of the New York Academy of Sciences*, 1124:161-180.
- Moll, J., R. de Oliveira-Souza, R. Zahn, and J. Grafman. 2008(b). The Cognitive Neuroscience of Moral Emotions. In: W. Sinnott-Armstrong (ed.) 2008:1-17.
- Moll, J., R. Zahn, R. de Oliveira-Souza, F. Krueger, and J. Grafman. 2005. The Neural Basis of Human Moral Cognition. *Nature Reviews Neuroscience*, 6:799-809.
- Osaka, N. (ed.) 2008. *Neural Correlates of Working Memory*. Kyoto University Press. (Japanese.)
- Paton, H.J. 1948. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. University of Chicago Press.
- Ribeiro, R.L., S. Pompéia, O. Francisco, and O.F.A. Bueno. 2005. Comparison of Brazilian and American Norms for the International Affective Picture System (IAPS). *Revista Brasileira Psiquiatria*, 27(3):208-215.
- Rozin, P., L. Lowery, S. Imada, and J. Haidt. 1999. The CAD Triad Hypothesis: A Mapping between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity). *The Journal of Personality and Social Psychology*, 76(4):574-586.

- Stratton-Lake, P. 2006. Moral Motivation in Kant. In: G. Bard. (ed.) 2006. *A Companion to Kant*. Blackwell, pp. 322-334.
- Tachibana, K. 2009a. Can Moral Enhancement be a Subject of Neuroethics? Neuroethics and the Governance of Brain Science. *Archive for Philosophy and the History of Science*, 11: 1-35. (Japanese.)
- . 2009b. On the Disquietedness of Moral Enhancement. *Enhancement, Society, and Human Nature*. UTCP Booklet, pp. 83-102. (Japanese.)
- Takahashi, H., M. Kato, M. Matsuura, D. Mobbs, T. Suhara, and Y. Okubo. 2009. When Your Gain Is My Pain and Your Pain Is My Gain: Neural Correlates of Envy and Schadenfreude. *Science*, 323: 937-939.
- Timmons, M. 2008. Toward a Sentimentalist Deontology. In: W. Sinnott-Armstrong (ed.) 2008: 93-104.
- Young, L., F. Cushman, M. Hauser, and R. Saxe. 2007. The Neural Basis of the Interaction between Theory of Mind and Moral Judgment. *Proceedings of the National Academy of Sciences*, 104 (20): 8235-8240.
- Valdesolo, P. and D. DeSteno. 2006. Manipulations of Emotional Context Shape Moral Judgment. *Psychological Science*, 17 (6): 476-477.
- . 2007. Moral Hypocrisy: Social groups and the Flexibility of Virtue. *Psychological Science*, 18 (8): 689-690.
- Zelazo, P.D., L. Qu, and U. Müller. 2005. Hot and Cool Aspects of Executive Function: Relations in Early Development. In: W. Schneider, R. Schumman-Hengsteler, and B. Sodian. (eds.) 2005. *Young Children's Cognitive Development*. Lawrence Erlbaum Associates, pp. 71-93.