AMAZING PAPERS IN NEUROSCIENCE
Does It Feel Right or Wrong? The Neuroscience of Moral Judgement

Ariane Delgado-Sánchez ‡
School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK, KY16 9JP; Present address: The University of Manchester, Manchester, United Kingdom, M6 8HD.

Moral judgement has been a topic of great interest through the history of philosophy and psychology, but the neural basis of this behavior remains elusive. Greene et al.'s (2001*)* paper is a pioneering one that opened doors to studying the neuroscience of moral judgement. Greene and colleagues used functional Magnetic Resonance Imaging (fMRI) to measure brain activity in humans as they grappled with moral decisions. The researchers found higher activation in brain areas associated with emotion when subjects were processing dilemmas in which an individual was directly hurt, and higher activation of areas associated with working memory when subjects processed dilemmas in which individuals were hurt as a consequence of an indirect action. The paper has received some criticism, but overall, is still quite relevant in the field. This work generated the field of moral neuroscience and has sparked further research and controversy. This paper's impact at the intersection of psychology and neuroscience and its engaging topic make it valuable for teaching. Furthermore, criticism of the paper also has pedagogic value as it can serve as a tool to promote students' critical evaluation skills. The presentation of this research would be best suited for Neuroscience and Psychology students being introduced to Cognitive Neuroscience.

How people make moral decisions has been a topic of general and scientific interest throughout human history. Initially this topic was studied by philosophers and psychologists. As new techniques developed, it began to be investigated from a neuroscience perspective. The pioneering work by Greene et al. (2001) is a representation of the first steps taken to understand the neural basis of moral judgment. Here, I review this paper and discuss its value as an educational tool.

The work of Greene et al. revolves around the presentation of two classical moral dilemmas: the "Trolley Dilemma" (Foot, 1968) and the "Footbridge Dilemma" (Thomson, 1976). In these two case scenarios the person reading them has to make a decision: sacrifice one person to save five people or do nothing and let five people die. Although the outcome is the same in both dilemmas, the procedure of sacrificing that one person is different. While in the Trolley Dilemma the sacrifice is made by hitting a switch, in the Footbridge Dilemma an individual has to be pushed off a bridge. This difference is impactful enough to make most people think that taking action is acceptable in the first dilemma and unacceptable in the second.

To test the neural bases of this phenomenon, the authors created a collection of dilemmas resembling the Trolley Dilemma (Impersonal Dilemmas) and the Footbridge Dilemma (Personal Dilemmas), as well as some Non-Moral Dilemmas. Dilemmas were first classified as moral or non-moral in a pilot study. Non-Moral Dilemmas were hypothetical situations similar to logic problems. Moral Dilemmas were grouped as personal or impersonal by two independent coders. To consider a dilemma personal the following criteria had to be met: *"actions (a) could reasonably be expected to lead to serious bodily harm (b) to a particular person or a member or members of a particular group of people (c) where this harm is not the result of deflecting an existing threat onto a different party"* (Greene et al., 2001). If these criteria were not met, a dilemma was assigned to the Moral Impersonal group.

After classification, all of the different dilemmas were presented to participants while they underwent a Functional Magnetic Resonance Imaging (fMRI) scan. fMRI measures the changes of blood oxygenation in the brain, which are associated with brain activity (reviewed in Buxton, 2013). In Greene et al. (2001), a different brain activation pattern was seen when subjects were presented with different types of dilemmas. Personal Dilemmas significantly correlated with a higher activation of areas that have been associated with emotion (medial frontal gyrus, posterior cingulate gyrus, and bilateral angular gyrus). Impersonal Dilemmas and Non-Moral Dilemmas significantly correlated with the activation of areas that have been related to working memory in previous research (middle frontal gyrus, bilateral parietal lobe).

The specific brain areas activated by the dilemmas are clearly seen in Figure 1 from the original article (Greene et al., 2001). These findings indicate that personal dilemmas evoke a greater 'emotional' processing of the situation, while impersonal dilemmas prompt a more 'cognitive' processing. However, as is common in cognitive neuroscience research, these conclusions must be analyzed with caution as they may be subject to a reverse inference effect. Reverse inference happens when a psychological function is inferred from a brain area activation (Poldrack, 2011). In the reviewed article there might be uncontrolled variables affecting the activation pattern, and therefore, conclusions cannot be definitive.

In their work, the researchers also discovered a second effect based on reaction time (RT) data. They observed that in the rare cases in which making a sacrifice in the Personal Dilemmas was considered acceptable by participants, the RT was longer. Furthermore, this decision was made despite high emotional brain activity. These results seemed to indicate that emotions interfered with the cognitive decision-making process. Later on, this conclusion served

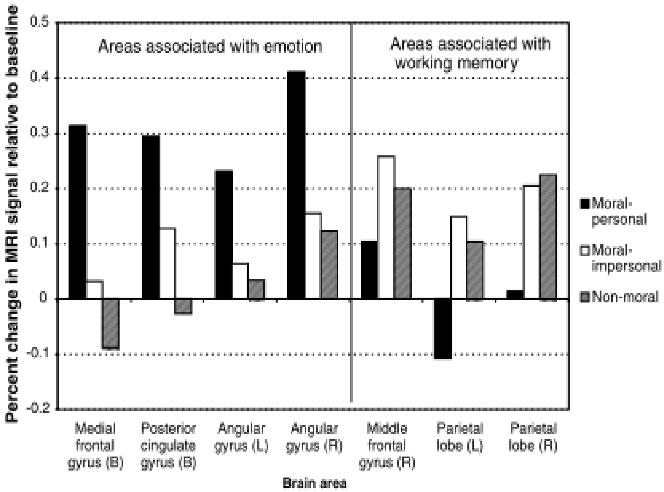

Figure 1. Effect of type of dilemma (personal/impersonal/non-moral) in brain activity (original figure from Greene et al., 2001).

as a first step for the development of the *Dual-Process Theory of Moral Judgment* (DPTMJ). This theory proposes the existence of two systems, one emotional and one cognitive/rational, that compete with each other in the process of moral decision making (Greene et al., 2004).

This work has generated a considerable amount of criticism, debate, and subsequent research that make it a perfect opportunity to instruct students about the process of science and develop their critical evaluation skills. The main source of debate surrounding this piece of work is centered on the pivotal conclusions drawn by the RT data. In 2009, McGuire et al. found that the RT effect was an artifact. These authors reanalyzed the original data, performing an item analysis in addition to the subject analysis performed in 2001. They discovered that the effect seen by Greene et al. (2001) was triggered by nine of the forty dilemmas presented to participants. Said dilemmas generated a very fast negative response in participants (more than 95% of participants immediately said that the action proposed was inappropriate). When the data was analyzed excluding said items the effects reported by Greene et al. (2001) disappeared.

Greene (2009) responded to the criticism by admitting that the RT effect found in the original experiment was an artifact. However, he defended his theory (the DPTMJ) stating that although the results from his study in 2001 had clear limitations, the results of his subsequent investigations supported it. For instance, he pointed to research in which they observed an interference in rational decision-making but not in emotional decision-making when cognitive load was increased. This supports the existence of the two competing systems (Greene et al., 2008).

The debate about how the brain processes information to make moral judgments is ongoing. In fact, this debate can be considered a part of a wider discussion on the topic of whether human cognition is ruled by multiple systems. Arguments in favour of both viewpoints have been made (e.g., Kahneman, 2011; Kruglanski & Gigerenzer, 2011) but no general agreement has been reached and research is still ongoing. Nevertheless, Greene's theory remains prevalent in the field. Since the seminal work by Greene et al. (2001), research has advanced and studies of the brain areas involved in the moral decision process have flourished. In this new research, the brain areas pointed out by Greene and colleagues (2001) continue to gather attention, although other areas have also been identified (reviewed in Pascual et al., 2013).

Importantly, the criticism of Greene's work has encouraged researchers to be more thorough when controlling external variables. For instance, new sets of moral dilemmas with better methodological controls have been created in an attempt to gather higher quality data (Lotto et al., 2014).

VALUE

Despite criticism, the work of Greene and has value, both from research and teaching perspectives.

From a research perspective, the investigation carried out by Greene et al. (2001) was ground-breaking. This paper represents the birth of the Neuroscience of Moral Judgment field, being a great example of how the development of new technologies such as fMRI can lead to new research. The work is highly referenced within the field and in textbooks (e.g., Gibbs, 2019; Ward, 2016) Furthermore, the theoretical framework (DPTMJ) that emerged from this investigation is still, despite criticism, one of the most relevant contemporary theories of Moral Judgment.

From a teaching viewpoint, the paper presents numerous advantages. Firstly, its clear and simple language makes it an easy paper to read even for those new to the topic. Secondly, it is a great first contact with the area of Moral Neuroscience as it gives students opportunity to learn how a new field of research emerges. Moreover, this work is also an engaging introduction to Cognitive Neuroscience in general. Morality is an issue that everyone can have an opinion without the need of specific knowledge in the area. This makes awakening student interest and generating debate easier. Moreover, the topic can also be a great opportunity for students to explore the relation between basic research and everyday life. Instructors could use this article to generate a discussion regarding the disciplines or areas of work in which basic scientific knowledge could be applied.

Finally, critiques of the article represent one of the greatest strengths of this work for teaching. Firstly, instructors can present to students the arguments by both McGuire et al. (2009) and Greene (2009), as well as the dilemmas that started the debate. This can form a foundation for debates in which students generate their own opinions. This would be an optimal way to help students develop critical thinking abilities while also promoting lively discussion. Furthermore, it gives instructors an opportunity to present neuroscience research as an ever-changing process that grows with new inputs and can be constantly improved and expanded. Many students may think of science as a linear process in which hypotheses are tested and turn out to be right or wrong. However, an article like this one gives students the opportunity to observe how science is in fact an iterative circular process in which hypotheses are continually reevaluated.

AUDIENCE

The reviewed paper could be a good tool in numerous teaching environments. It would be best suited for 2nd, 3rd or 4th year Neuroscience or Psychology students. Although the content is easy to understand without prior knowledge of the specific area, having basic understanding of psychology and neuroscience is beneficial. In terms of specific classes, this paper fits well in an introductory class on Cognitive Neuroscience or an advanced class on Human Cognitive Development. Furthermore, it could also be used in more advanced research methodology classes to treat issues such as reverse inference, artifacts and reanalysis of data.

Overall, this paper brings together the disciplines of Psychology, Neuroscience and Ethics. It is therefore an amazing paper for engaging students across multiple disciplines.

REFERENCES

Buxton RB (2013) The physics of functional magnetic resonance imaging (fMRI). Reports Prog Phys 76:096601.
Foot P (1968) The Problem of Abortion and the Doctrine of the Double Effect. Oxford Rev:1–7.
Gibbs JC (2019) Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt., 4th edition New York, NY: Oxford University Press.
Greene JD (2009) Dual-Process Morality and the Personal / Impersonal Distinction : A Reply to McGuire, Langdon, Coltheart, and Mackenzie. J Exp Soc Psychol 45:581–584.
Greene JD, Morelli SA, Lowenberg K, Nystrom LE, Cohen JD (2008) Cognitive load selectively interferes with utilitarian moral judgment. Cognition 107:1144–1154.
Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The Neural Bases of Cognitive Conflict and Control in Moral Judgment. Neuron 44:389–400.
Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI Investigation of Emotional Engagement in Moral Judgment. Science 293:2105–2108.
Kahneman D (2011). Thinking, Fast and Slow. London: Penguin Books.
Kruglanski AW, Gigerenzer G (2011) Intuitive and deliberate judgments are based on common principles. Psychological Review 118: 97–109.
Lotto L, Manfrinati A, Sarlo M (2014) A New Set of Moral Dilemmas : Norms for Moral Acceptability, Decision Times, and Emotional Salience. J Behav Decis Mak 27:57–65.
McGuire J, Langdon R, Coltheart M, Mackenzie C (2009) A reanalysis of the personal / impersonal distinction in moral psychology research. J Exp Soc Psychol 45:577–580.
Pascual L, Rodrigues P, Gallardo-pujol D, Shepherd GM (2013) How does morality work in the brain? A functional and structural perspective of moral behavior. Front Integr Neurosci 7:1–8.
Poldrack RA (2011) Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. Neuron 72:692–697.
Thomson J (1976) Killing, Letting Die and the Trolley Problem. Monist 59:204–217.
Ward J (2016) The student's guide to social neuroscience., 2nd edition New York, NY: : Psychology Press.

Received April 4, 2020; revised April 24, 2020; accepted May 24, 2020.

Address correspondence to: Ariane Delgado-Sanchez, Human Pain Research Group, School of Biological Sciences, University of Manchester, Clinical Sciences Building, Salford Royal NHS Foundation Trust, Manchester, England, , United Kingdom, M6 8HD. Email:ariane.delgadosanchez@manchester.ac.uk

www.funjournal.org