

# Nichols: the Catastrophe Case and How We Make Moral Judgments on Killing Innocent People

## Nichols: o Caso Catástrofe e Como Fazemos Juízos Morais sobre Matar Pessoas Inocentes

CINARA NAHRA\*

**Abstract:** The aim of this article is to discuss Nichols and Mallon's article "moral dilemmas and moral rules" and what I call their three thesis, i.e: THESIS 1- People are not absolutist deontologists - THESIS 2: People appreciate a distinction between weak and all-in impermissibility - THESIS 3: There are two partly independent mechanisms underlying moral judgment, one based on rules and the other based on consequences of the actions. Based on these three theses I discuss further the possibility of establishing a model able to explain people's judgments in relation to moral dilemmas.

**Keywords:** Bioethics. Deontological Ethics. Moral Dilemmas.

**Resumo:** O objetivo deste artigo é discutir o artigo de Nichols e Mallon "Dilemas morais e regras morais" e o que eu chamo de suas três teses, a saber: Tese 1 - As pessoas não são absolutamente deontológicas em relação a julgamentos morais relativos a tirar a vida; Tese 2 - As pessoas apreciam uma distinção entre inadmissibilidade fraca e forte; Tese 3 - Existem dois mecanismos parcialmente independentes por trás de nossos julgamentos morais, um baseado em regras e o outro baseado nas consequências de nossas ações. A partir destas três teses discuto posteriormente a possibilidade de estabelecer um modelo capaz de explicar os julgamentos das pessoas em relação aos dilemas morais.

**Palavras-Chave:** Bioética. Ética Deontológica. Dilemas morais.

---

\* Doutora em Filosofia pela University of Essex, Inglaterra. Professora do Departamento de Filosofia da Universidade Federal do Rio Grande do Norte (UFRN). cinaranahra@hotmail.com.

## Introduction

Nichols and Mallon<sup>1</sup> propose an important distinction between what they call judgments of “weak impermissibility” and judgments of “all-in impermissibility”. For them judgments that an action violated a rule are called judgments of “weak impermissibility”, while judgments that an action was wrong, all things considered, will be called judgments of “all-in impermissibility”. To show this they put forward the catastrophe dilemma, which is the following<sup>2</sup>:

A train is transporting an extremely dangerous artificially produced virus to a safe disposal site. The virus is profoundly contagious and if the virus were to be released into the atmosphere, billions of people would die from it. Jonas is one of the scientists who was responsible for ensuring that the virus would be destroyed and he is watching the train from a footbridge. As the train approaches, he sees through his binoculars that there is a powerful bomb planted on the tracks ahead, and there is no way for him to communicate with the train operators to get them to stop the train in time. If the train passes over the bomb, it will explode and the virus will be released into the environment with catastrophic consequences. There is a large stranger looking over the footbridge next to Jonas. Jonas knows that the stranger has nothing to do with the bomb, but the only way to stop the train from hitting the bomb is to push this stranger over the railing. For unlike Jonas’s body, the stranger’s body is big enough that it will bring the train to a halt, although this will kill the stranger. Jonas proceeds to push the stranger over the railing, which kills the stranger, but it prevents the explosion and saves billions of people from dying from the virus.

Nichols and Mallon then carried out a research, asking the participants to answer two questions. a) Did Jonas break a moral rule by pushing the stranger over the railing? b) All things considered, did Jonas do the wrong thing? Nichols and Mallon then reported that 68% of participants

---

<sup>1</sup> Schaun, Nichols; Mallon, Ron “Moral dilemmas and Moral rules” *Cognition* 100 (2006) p. 530–542

<sup>2</sup> *Ibid.*, p.538.

said the actor broke a moral rule, but only 24% said the action was, all things considered, the wrong thing to do.

The authors interpret people's answer in this catastrophe case as a proof that when the consequences are overwhelmingly bad, the consequences can trump the moral rules. The results of this experiment are especially interesting because it is well known that in the standard cases of the "footbridge fat man", the ones where there is not such a high number of people dying,<sup>3</sup> the majority of people typically answer that it is wrong to push the man over the bridge.

Greene and all reported<sup>4</sup> data from a research carried out for them. In this research 21% of the participants answered that "it is appropriate to push the stranger onto the tracks in order to save five workmen". Notice that there is here an inversion of percentage. While in the typical case of the footbridge dilemma only 21% of the people said that it is appropriate to push the man, in Nichols experiment, where billions could die, only 24% said that the action was, all things considered, wrong. Why is this? The more plausible answer, in line with what Nichols concludes, is that people recognize that a rule was broken, but, given the circumstances, i.e., given the possibility of billions of people dying if the rule were not broken, they consider that the action was, all things considered, not wrong.

It seems, actually that there are three thesis in this article of Nichols and Mall:

1) People are not absolutist deontologists (people think that sometimes it is all-in permissible to do something that violates a moral rule, even the rule that forbids killing innocent people);

---

<sup>3</sup> On the dilemma of the *fat man* see Thomson, Judith "The Trolley Problem" *The Yale Law Journal*, Vol. 94, No. 6 (May, 1985), pp. 1395-1415 and Foot, P 'The Problem of Abortion and the Doctrine of the Double Effect', in *Virtues and Vices and other essays in moral philosophy*, Oxford Review, n. 5 (1967).

<sup>4</sup> Greene, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., Cohen, J.D. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, Vol. 107, 1144-1154 (2008).  
Supplementary materials at <http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-CogLoadSupMats.pdf> (Accessed 17/9/2014).

2) People appreciate a distinction between weak and all-in impermissibility.

3) There are two partly independent mechanisms underlying moral judgment. On one hand, people have a general capacity to reason about how to minimize bad outcomes, but on the other hand, people have a body of rules prescribing certain actions and this body of rules cannot be subsumed under the capacity to reason about how to minimize bad outcomes.

I will now elaborate on each one of these thesis and propose some discussions from them.

### **THESIS 1 – People are not absolutist deontologists**

This means that when people make moral judgments they do not base their judgments only on moral rules. It means that people are willing, in some circumstances, to break moral rules for some reason, or a set of reasons, as, for example, when the consequences of abiding by a deontological prohibition might be disastrous. This is what the experiment on the catastrophe case seems to be suggesting.

However, here I want to propose a reflection. We know from experiments that overriding deontological rules do not happen only in catastrophe cases. Typically, the majority of people also say that they would push a button to divert a train in order to kill one person, instead of five. The fact that people in general think that it is right to push the button and divert the train, shows already that we are not absolutist deontologists.<sup>5</sup> So, what is the novelty in the catastrophe experiment?

One of the novelties, I think, is that the catastrophe dilemma puts more wood on the dilemmas fire. Why do people admit pushing someone to their death in order to save 1 billion fellow human beings but they would not admit to do the same in order to save 5 people? An important subject that emerges from Nichols experiment is not only the awareness that we are not absolutist deontologists, but also a discussion on if it is possible to describe certain general conditions that make people switch from their deontological

---

<sup>5</sup> See for example Hauser, *Marc Moral Minds: How Nature designed our universal sense of right and wrong* (USA: Harper Collins Publishers, 2006).

judgments for a utilitarian one and, if this is the case, “what are the conditions that make people not to base their judgments in a deontological way anymore”?

**THESIS 2: People appreciate a distinction between weak and all-in impermissibility.**

According to Nichols judgments of weak impermissibility are those where people recognize that a moral rule was broken, and judgment of all-in impermissibility are those where people recognize that despite the fact that a moral rule was broken, the action was, all things considered, wrong. It means, confirming thesis 1, that people are not absolutist deontologists, but it also poses the question about what is involved in “all in impermissibility” judgments, why do people say that it is “okay” to do something despite the fact that it breaks a moral rule? Nichols proposes that assessment of all-in impermissibility implicates three factors: a) cost/benefit analysis, b) checking for rule violations, and c) emotional activations.

An analysis of cost/benefit with costs outweighing benefits can sway people not to give moral rules the final word when they make moral judgments, judging that certain actions are permissible despite the fact that they break a moral rule. The example used by the authors in the experiment 3 of their article is the moral rule that forbids killing innocent people. The experiment suggests that people admit and accept the rule but admit to breaking it for utilitarian reasons, in order to save more lives.

My question is then: Wouldn't it be possible to rewrite the rule in order to incorporate the utilitarian element? Wouldn't it be possible to say that when people answer that it is not all-in impermissible to kill innocent people in order to avoid the death of billions, they are actually following a rule that incorporates and admits to a utilitarian calculation? The rule, actually, could be rewritten so “it is forbidden to kill innocent people in a personal way, unless killing those innocent people will save a much greater number of people”.

In a recent article Nichols and Mallon<sup>6</sup> in questioning the dual process theory, which suggests that moral judgment will be either unconsciously generated intuitions or consciously available but effortful reasoning, raise another possibility, which is, that there are “rules that are consciously available and effortlessly applied in moral judgments”. The example they give of such rules is the example of incest, but why the rule that it “is forbidden to kill innocent people, unless it will save a much higher number of lives” wouldn’t also be one of these rules?

If this is the case we still would have to explain why this rule is not applied in the mainstream case of the fat man on the bridge, and then the answer could be precisely that the emotional activation counteracts the rule, or maybe, it’s because there is another rule counteracting this rule, as for example the rule that “it is forbidden to kill someone in a personal way”. Therefore, we could speculate that the difference between the bridge dilemma and the catastrophe case is precisely the benefit rate. While in the first case only four people are saved, in the second case billions of people are, which could suggest that quantitative considerations are relevant in moral judgments and can be the turning point from deontological to utilitarian judgments.

**THESIS 3: There are two partly independent mechanisms underlying moral judgment, one based on rules and the other based on consequences of the actions.**

What we have discussed above corroborates thesis 3 of Nichols paper, which sustains that there are two independent mechanisms underlying moral judgments, one that is related to the minimization of bad outcomes and the other related to a body of rules. Nichols suggestion, in my view, corroborates the view<sup>7</sup> that there are two mechanisms working in moral judgment, one primarily deontological (the one based on rules and absolute

---

<sup>6</sup> Mallon, Ron and Nichols, Shaun “Dual Processes and Moral Rules” *Emotion Review* 3: 284 (2011)

<sup>7</sup> This view that there are two mechanisms working in moral judgements, one deontological and the other one utilitarianist came specially from the works of Joshua Greene and what he calls the dual process theory of moral judgement.

prohibitions) and another primarily utilitarian (based on cost/benefit reasoning). Nichols seems to suggest as well, from thesis 2, that people recognize moral rules and follow them, up to the moment that, for utilitarian reasons, they admit to breaking them.

If my interpretation is right it seems to me that the key question to be answered would be to determine what makes people switch from one way of judging to another. Let's look again at Nichols words<sup>8</sup>:

The philosophical project is to consider our intuitions about a wide range of dilemmas and to determine a set of principles that captures our intuitions about the cases (...). One goal of the philosophical investigations has been to develop a unified normative theory that will accommodate our intuitions about such cases. This goal has been exceedingly difficult to meet, and few would maintain that philosophers have succeeded in finding a unified normative theory that fits with all of our intuitions about moral dilemmas. Although philosophers have not produced a unified normative theory that accommodates all of our intuitions about moral dilemmas, the empirical work has reinvigorated the investigation of moral dilemmas.

I have no doubt that Nichols work gives us some good insights in the direction of a further normative theory able to explain people's judgments in relation to moral dilemmas. The empirical work that he has carried out, seems to me to be pointing to an explanatory model of moral judgments which concedes that for an action to be considered impermissible a rule must exist forbidding actions of this type, for example a deontological rule which forbids to kill innocent people (you shall not kill!).

However, as Nichols adverts, a violation of a rule and emotional activation does not necessitate the judgment that the action is wrong, because people could consider, all things considered, that the action is the right thing to do, as it happens in catastrophe cases. Nevertheless, what are these "all things considered"? They are above all the cost/benefit analyses, which makes, for example, people say that it is right to push an innocent person onto a railway track in order to save billions of lives. But is this the

---

<sup>8</sup> Shaun, Nichols; Mallon, Ron "Moral dilemmas and Moral rules" Op.cit., p. 531.

end of the story, i.e., the majority of people would under any circumstances accept that it is right to kill an innocent in order to save billions?

Here, let us return to the classical philosophical discussion with Philippa Foot's example of the bad man and Bernard Williams's example of Jim<sup>9</sup>. Greene has shed some light on this discussion carrying out empirical experiment that corroborate the intuitions of these philosophers. Greene tests what he calls the *modified safari dilemma*<sup>10</sup> where a group of terrorists promises to save your life and the lives of the children, if you personally kill one of the hostages who is being held with you. The percentage of utilitarian answers in the *modified safari* are only 22%, according to Green's figures, i.e., only 22% said that it is appropriate to kill one of your fellow hostages in order to save the others which suggests that the cost-benefit utilitarian reasoning can still be overcome by deontological considerations.

Would the people's answer be the same if instead of only five people's the lives of billions were at risk? Only experiments could give us this answer, but I think that a pattern is starting to emerge here, a pattern that I call deontoutilitarianist<sup>11</sup>. At least in those questions related to dilemmas involving killing others, people seem to accept the deontological role that we should not kill. They are willing, however, to accept exceptions to this rule for utilitarian reasons, in order to save more lives. They become deontological again, nevertheless, if the killing involves some kind of physical contact or proximity with the person who will be killed.

But again, for utilitarian reasons, such as a huge number of lives being saved, people admit to breaking the rule. Finally, these people probably become deontological again if the killing has to be done to satisfy, for example, the outrageous requirements of a perceived evil person who

---

<sup>9</sup> See Foot, Philippa op.,cit 'The Problem of Abortion and the Doctrine of the Double Effect', in *Virtues and Vices and other essays in moral philosophy*, Oxford Review, n. 5 (1967) and for Bernard Williams example of Jim see Williams, B 'A Critique of Utilitarianism' in B.Williams, Smart, C *Utilitarianism For and Against* (Cambridge:Cambridge University Press:1973).

<sup>10</sup> Greene and others Op. Cit. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, Vol. 107, 1144-1154 (2008). Supplementary materials at <http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-CogLoadSupMats.pdf> (Accessed 17/9/2014).

<sup>11</sup> See Nahra, Cinara "Our deontological-utilitarian (deontoutilitarian) minds" *Filosofia Unisinos* v. 14 n.2 (2013) p.139-151.



blackmails you, threatening to kill more people if you refuse to comply and demands that you actively carry out the killing. Could this be an emerging pattern in the way that the majority of people make their moral judgments in dilemmas involving killing in order to save others? To answer this question it is time to follow in Nichols steps, get out of the chair and experiment, experiment, experiment.

## **Bibliography**

FOOT, Phillipa 'The Problem of Abortion and the Doctrine of the Double Effect', in *Virtues and Vices and other essays in moral philosophy*, Oxford Review, n. 5 (1967).

GREENE, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., Cohen, J.D. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, Vol. 107, 1144-1154 (2008).

GREENE, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., Cohen, J.D. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, Vol. 107, 1144-1154 (2008).

Supplementary materials at  
<http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-CogLoadSupMats.pdf> (Accessed 17/9/2014).

HAUSER, Marc *Moral Minds: How Nature designed our universal sense of right and wrong* (USA: Harper Collins Publishers, 2006).

MALLON, Ron and NICHOLS, Shaun "Dual Processes and Moral Rules" *Emotion Review* 3: 284 (2011).

WILLIAMS, B 'A Critique of Utilitarianism' in B.Williams, *Smart, C Utilitarianism For and Against* (Cambridge: Cambridge University Press:1973).

NAHRA, Cinara "Our deontological-utilitarian (deontoutilitarian) minds" *Filosofia Unisinos* v. 14 n.2 (2013) p.139-151.

SCHAUN, Nichols; Mallon, Ron "Moral dilemmas and Moral rules" *Cognition* 100 (2006).

THOMSON, Judith "The Trolley Problem" *The Yale Law Journal*, Vol. 94, No. 6 (May, 1985).

**Endereço Postal**

PUCRS-Programa de Pós-Graduação em Filosofia

Av. Ipiranga, 6681 – Prédio 5 – Sala 608

CEP: 90619-900 – Porto Alegre – RS

**Data de Recebimento**

17 de setembro de 2014

**Data de Aceite para Publicação**

15 de dezembro de 2014