Neuroethics (2018) 11:195–209
https://doi.org/10.1007/s12152-018-9362-y

ORIGINAL PAPER

On the normative insignificance of neuroscience and dual-process theory

Peter Königs

Received: 7 January 2018 / Accepted: 27 March 2018 / Published online: 4 April 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract According to the dual-process account of moral judgment, deontological and utilitarian judgments stem from two different cognitive systems. Deontological judgments are effortless, intuitive and emotion-driven, whereas utilitarian judgments are effortful, reasoned and dispassionate. The most notable evidence for dual-process theory comes from neuroimaging studies by Joshua Greene and colleagues. Greene has suggested that these empirical findings undermine deontology and support utilitarianism. It has been pointed out, however, that the most promising interpretation of his argument does not make use of the empirical findings. In this paper, I engage with recent attempts by Greene to vindicate the moral significance of dual-process theory and the supporting neuroscientific findings. I consider their potential moral significance with regard to three aspects of Greene's case against deontology: the argument from morally irrelevant factors, the functionalist argument and the argument from confabulation. I conclude that Greene fails to demonstrate how neuroscience and dual-process theory in general can advance moral theorizing.

Keywords Debunking arguments · Joshua Greene · Deontology · Utilitarianism · Neuroscience · Dual-Process Theory, Selim Berker

P. Königs (✉)
Department of Philosophy, University of Tübingen, Bursagasse 1, 72070 Tübingen, Germany
e-mail: peter.koenigs@uni-tuebingen.de

Introduction

According to the dual-process account of moral judgment, deontological and utilitarian judgments typically stem from two different cognitive subsystems. Deontological responses are effortless, intuitive and emotion-driven whereas utilitarian responses are effortful, reasoned and dispassionate. The philosopher and psychologist Joshua Greene has not only pioneered dual-process theory with two seminal neuroimaging studies on people's responses to moral dilemmas. He has also suggested that dual-process theory, and thus indirectly his neuroscientific findings supporting it, have substantive implications for normative moral philosophy. More specifically, Greene has argued that dual-process theory casts doubt on deontology and supports utilitarianism. Dual-process theory is thus provided as an example of how the cognitive science of ethics can advance normative theorizing.

While this sounds promising, the devil is, as always, in the details. The empirical work done by Greene and others has certainly shed new light on this old normative controversy, but whether and exactly how it may contribute to its resolution is less clear. In fact, it is unclear whether Greene's argument for utilitarianism even requires the neuroscientific findings whose normative significance he has so vocally professed. Selim Berker has forcefully argued that the most promising interpretation of Greene's argument renders the neuroscientific findings – and dual-process theory in general – superfluous.

Greene's empirically informed argument for utilitarianism has since undergone several modifications, and

he has gone to some lengths to demonstrate the normative relevance of his empirical work.[1] The purpose of this article is to re-evaluate the moral significance of the dual-process account and the underlying neuroscientific findings. I will confirm and expand upon Berker's criticism by engaging with Greene's more recent writings. While dual-process theory and Greene's neuroscientific work may of course still turn out to be normatively relevant, Greene's attempts to demonstrate their significance are so far unconvincing.[2]

The article will be structured as follows: I will first, in section 2, provide a more detailed account of Greene's argument and Berker's critique and say a word on the dialectic of this paper. In sections 3, 4 and 5, I discuss three possible ways of vindicating the normative significance of dual-process theory and explain why none of them is convincing. They relate to three different versions or components of Greene's argument: the argument from morally irrelevant factors (section 3), the functionalist argument (section 4) and the argument from confabulation (section 5). I conclude, in section 6, with some brief comments on Greene's emphasis on the normative significance of empirical findings other than those directly related to dual-process theory.

Greene's argument and its problems

Greene and colleagues' pathbreaking fMRI studies are the founding studies of dual-process theory. Their first fMRI study explained why people respond differently to similar sacrificial dilemmas, such as different versions of the notorious trolley dilemma. The SWITCH scenario, for instance, which involves hitting a switch to divert the trolley onto a side track with only one rather than five workers on it, typically elicits the utilitarian response that it is appropriate to hit the switch. The FOOTBRIDGE dilemma, by contrast, which involves shoving a heavy person from a footbridge onto the track below to block the trolley and save the five workers, tends to elicit the deontological intuition that it is not appropriate to sacrifice the heavy person. Greene and colleagues found that FOOTBRIDGE-like dilemmas, characterized by up close and personal violations, engage brain areas associated with emotional processing, whereas SWITCH-like dilemmas, involving impersonal violations, lead to increased activity of brain areas associated with working memory.[3] In a follow-up study, Greene and colleagues showed that high-reaction-time personal dilemmas are associated with increased activity of brain regions linked to cognitive conflict and of brain regions linked to abstract reasoning as compared to low-reaction-time personal dilemmas. This supports the hypothesis that long reaction times in response to personal dilemmas are due to a conflict between a prepotent emotional deontological response and a consequentialist reasoning process. It was also found that consequentialist responses to high-reaction-time personal dilemmas are correlated with higher activity of the 'cognitive' brain areas.[4]

Besides these two neuroimaging studies, there is now a plethora of neuroscientific and other research that points in a similar direction.[5] Therefore, while Greene and colleagues' fMRI studies provide the most prominent evidence for dual-process theory, Greene is right to point out that the validity of this theory does not strictly depend on these studies.[6]

Dual-process theory is associated with a hypothesis concerning the genealogy of deontological intuitions. The fact that deontological responses originate from the intuitive and emotional subsystem indicates that they are the products of natural selection. Greene explains that "when Nature needs to get a behavioral job done, it does it with intuition and emotion wherever it can."[7] The fact that our intuitions vary as a function of whether or not harm is inflicted in an up close and personal manner also suggests an evolutionary explanation. It is likely that we have evolved an innate moral aversion to interpersonal violence. But it is to be expected that only the infliction of personal harm pushes our moral buttons, since our ancestral environment did not include the technical means to inflict harm in impersonal ways.[8]

Based on these empirical results, Greene has developed a so-called debunking argument against our

[1] Especially in [1–4].

[2] I am (like Berker and Greene) specifically considering the significance of the neuroscientific findings that support dual-process theory, rather than that of neuroscience in general.

[3] Greene et al. [5]. The study also tested reaction times, but the data were misinterpreted [6, 7].

[4] Greene et al. [8].

[5] For an overview of the evidence, see Greene [3].

[6] Greene [1], pp. 8, 20; [3], pp. 705–706. This renders criticism specifically of the neuroimaging studies less damaging (e.g. [9]).

[7] Greene [10], p. 60, see also pp. 70–71.

[8] Greene [11], p. 345; [12], p. 59; [10], p. 43; Greene et al. [8], pp. 389–390. Similar evolutionary debunking arguments are suggested for our moral condemnation of incest and our retributive intuitions [10].

deontological intuitions. Debunking arguments are arguments that undermine a belief or doctrine by showing its causal origins to provide an undercutting defeater.[9] Greene has claimed that we should distrust our deontological intuitions on the grounds of what we know about their psychological underpinnings. The original statement of the argument can be interpreted in two different ways.[10]

According to the first interpretation, which yields the argument from evolutionary history, we should distrust our deontological gut reactions because they are the products of natural selection. Since natural selection is not a truth-tracking process, our deontological intuitions have no evidential force. The problem with this argument is that it is unlikely to confer a competitive advantage to utilitarians. As many have pointed out, utilitarian intuitions must be expected to fall victim to genealogical debunking explanations, too.[11] It is therefore unlikely that the evolutionary debunking project can be contained in such a way that utilitarian principles survive, and Greene appears to concur.[12]

According to the other interpretation, which gives us the argument from morally irrelevant factors, we should distrust our deontological intuitions because they are responsive to factors that we know to be morally irrelevant. Whether harm is inflicted in a personal or impersonal fashion is irrelevant from the moral point of view. As Greene jokingly observes, "[w]ere a friend to call you from a set of trolley tracks seeking moral advice, you would probably not say, 'Well, that depends. Would you have to *push* the guy, or could you do it with a switch?'"[13] The idea is that if it is morally permissible to sacrifice one person to save five in an impersonal way and if 'personalness' is not a morally relevant consideration, then it should also be permissible to sacrifice one person to save five in a personal way. The argument is now informed by findings from a 2009 study by Greene and colleagues that was specifically designed to provide a more precise account of which factors trigger deontological responses. It found that our responses are sensitive to 'personalness' in that they are sensitive to whether or not the "force that directly impacts the other is generated by the agent's muscles, as when one pushes another with one's hands or with a rigid object" (but only if the violation is intended rather than merely foreseen).[14]

Selim Berker, however, has pointed out that the argument from morally irrelevant factors does not rely on Greene and colleagues' neuroscientific findings and the dual-process theory that these findings support. Rather, the argument is based on assumptions about which factors trigger deontological responses and on normative assumptions about whether these factors are morally relevant. The psychological nature of deontological and utilitarian responses – whether they are emotion-driven or not, whether they stem from the same or different cognitive systems, etc. – is entirely beside the point. Thus, given the above-noted problem with the argument from evolutionary history, it appears that the most promising version of Greene's debunking of deontological intuitions renders dual-process theory and the neuroscientific findings normatively insignificant.[15]

This is the challenge that Greene has recently attempted to defuse. He has argued against Berker that dual-process theory *does* support the argument from morally irrelevant factors, and he has proposed yet another argument for utilitarianism that is informed by dual-process theory. I have dubbed the latter the functionalist argument. In a nutshell, the functionalist argument states that we should favor utilitarian over deontological solutions to unfamiliar moral problems because our automatic, deontological intuitions have not evolved to deal with such problems. As noted above, I will also examine a third way in which Greene might defend the normative significance of the neuroscientific results: Greene speculates that more elaborate deontological theories, which do not invoke the debunked

[9] Kahane [13], p. 106.

[10] These are at least the two most natural and charitable interpretations of the argument put forth in Greene [10]. For further possible interpretations, see Berker [14]. For another general discussion of Greene's argument, see Sauer [15]. A similar argument, based on the same empirical findings, was developed by Singer ([16]).

[11] Berker [14], p. 319; Kahane [13, 17]; Mason [18]; Tersman [19]. Note that Greene acknowledges that utilitarianism depends on intuitions, too, but he takes these intuitions to be about general principles rather than particular actions and to differ psychologically from deontological intuitions ([1], pp. 19–20; [3], p. 724).

[12] At least, he has not attempted to defend the argument from evolutionary history, and he has more recently expressed doubts concerning the evolutionary hypothesis ([4], p. 68).

[13] Greene [20], p. 176.

[14] Greene [6], p. 365. By contrast, a dilemma was originally classified as 'personal' if "the action in question (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." ([5], p. 2107). While our responses are sensitive to the conjunction of two factors (personal force plus intention), the argument from morally irrelevant factors focuses primarily on the irrelevance of the former factor.

[15] Berker [14].

deontological intuitions, are the results of confabulatory post hoc rationalization. This argument from confabulation is an important complement to Greene's attack on deontological intuitions, and Greene's presentation of it makes it appear to rely on the neuroscientific findings.

I will argue that none of these three approaches are successful. If my diagnosis is correct, Greene has failed to demonstrate how his neuroscientific research and dual-process theory in general may advance normative theorizing. I am assuming that Greene must show two things to establish their normative significance. First, the empirical findings must play a role in the argument. Second, the argument must actually be sound.[16] The three attempts I consider fail to meet these criteria in different ways: The argument from morally irrelevant factors does not depend on dual-process theory, thus failing to meet the first criterion. Greene then puts forth an argument that complements the argument from morally irrelevant factors, which is supposed to correct a weakness of this argument. But this complementary argument, which does invoke dual-process theory, is not sound, thus failing to meet the second criterion. The functionalist argument does make use of dual-process theory, but it is intrinsically unconvincing. It, too, fails to meet the second criterion. And, contrary to appearance, the argument from confabulation does not depend on the neuroscientific findings, thus failing to meet the first criterion.

The argument from morally irrelevant factors

In response to Berker's criticism, Greene contends that dual-process theory is normatively significant as it gives us reason to believe that consequentialist responses are not similarly vulnerable to the argument from morally irrelevant factors. Greene thereby addresses two problems at once: He demonstrates the moral significance of dual-process theory (first problem) by arguing that it explains why a similar argument against consequentialist judgments is unlikely to succeed (second problem). This brief summary of what I take to be Greene's reasoning will require some unpacking.[17]

But before I examine Greene's argument in more detail, let us briefly dwell on this second problem, also raised by Berker, that utilitarian judgments might be vulnerable to a parallel challenge. Berker observes that "it is open to the defender of deontology to reply that, intuitively, the faculty eliciting consequentialist reactions is also responding to morally irrelevant factors, or failing to respond to morally relevant ones."[18] Note that this parallel anti-consequentialist argument would probably take the negative form, stating that utilitarian judgments *fail* to be sensitive to factors that *are* relevant. Here is just one example of how such a parallel argument might go: Consider CHILD, which is a variation of the traditional SWITCH dilemma except that the person on the side track is your own child. Saving the five workers would require killing your own child. A critic of utilitarianism could plausibly argue that there *is* a morally relevant difference between SWITCH and CHILD, namely that the victim in CHILD is your own child. This is why the consequentialist response to CHILD fails to be sensitive to a morally relevant factor. Examples like this can easily be multiplied, as there are many factors apart from those bearing on the maximization of welfare that are widely felt to be normatively relevant.[19] And for every such factor we could formulate a parallel argument against utilitarian judgments.

This objection does not imply that Greene's attack on specific deontological intuitions fails. It does not vitiate the finding that certain deontological intuitions are responsive to a morally insignificant factor (personal force). What it does mean, though, is that Greene will have a hard time establishing utilitarianism by looking at which judgments are sensitive to morally (ir)relevant factors. While he may be able to debunk some deontological intuitions as responsive to irrelevant factors, it is at least doubtful that this approach will eventually vindicate utilitarianism.

Victor Kumar and Richmond Campbell have come to the rescue of Greene by drawing attention to an asymmetry between Greene's anti-deontological argument and parallel anti-consequentialist arguments.[20] They observe that the latter are less effective because they rely on normative premises that are controversial, whereas Greene's argument invokes an intuition that even deontologists accept. No deontologist would want to say that

[16] I am here adopting the dialectic of Berker's argument, who contends that Greene's arguments either "rely on a shoddy inference" or on premises "that render the neuroscientific results irrelevant to the overall argument."([14], p. 294)

[17] It is probably fair to say that Greene's notes on this issue, which are labelled as 'work in progress', are somewhat sketchy. Below, I present what I hope is their most charitable interpretation.

[18] Berker [14], p. 325.

[19] Berker [14], p. 325.

[20] Kumar and Campbell [21], pp. 314–15.

personal force makes a moral difference.[21] By contrast, utilitarians are (relatively) happy to assert that kinship – or indeed *any* factor that does not affect the maximization of welfare − is morally irrelevant. The anti-consequentialist argument thus threatens to beg the question against the consequentialist. Greene's argument, by contrast, is convincing even to deontologists. Kumar and Campbell have, I think, correctly identified an important strength of Greene's argument. It also means that Berker's claim that Greene's argument fails to "advance the dialectic on the relative merits of deontology versus consequentialism" is unfair.[22] It does advance the debate precisely because it rests on an assumption that even deontologists cannot dispute. At the same time, however, this asymmetry does not entirely defuse Berker's challenge. It does not mean that these parallel anti-consequentialist arguments are altogether without force so that Greene need not be concerned about them. Surely, the fact that a great many people intuit that utilitarian judgments fail to respond to morally relevant factors is a serious problem for utilitarians. An objection along these lines might not be particularly original and thus do less to 'advance the dialectic', but it cannot simply be dismissed, either. Besides, the fact that Greene makes considerable efforts to respond to this challenge indicates that he, too, takes it seriously.

This being said, let us consider Greene's attempt to defuse this problem by appeal to dual-process theory. Greene writes: "Why not suppose, as Berker does [...], that consequentialist 'intuitions' are as much to blame as deontological ones? The answer is that there is a deep cognitive asymmetry between consequentialist and deontological thinking, as posited by the dual-process theory."[23] He then goes on to explain that consequentialist responses are reasoned in that they involve the conscious application of a moral principle (that of maximizing welfare). By contrast, deontological intuitions are automatic and emotional, and when people give deontological responses, they are often unaware of the principles that govern their responses (e.g. the doctrine of double effect).[24] Indeed, consequentialist intuitions have been shown to be psychologically so unlike ordinary intuitions that they do not even qualify as intuitions in the psychological sense of the term. They are only intuitions in the philosopher's sense. And because they are so different, they are less likely to be vulnerable to the argument from morally irrelevant factors, as it casts doubt specifically on 'psychological' intuitions:

> "In short, characteristically consequentialist judgments are not intuitive in the psychological sense, but characteristically deontological judgments are. [...] More generally, our mysteriously variable moral intuitions are a nuisance for consequentialists, but they are [...] the lifeblood of deontological theorizing. For these reasons, evidence that our intuitions [in the psychological sense] are unreliable is a point in favor of consequentialism and a point against deontology."[25]

Greene and colleagues' neuroscientific results matter for this argument, because they support the dual-process theory, which explains this difference between deontological and consequentialist responses.[26] That is, Greene is not saying that dual-process theory plays a role in showing that deontological responses are sensitive to irrelevant factors (here apparently concurring with Berker). But dual-process theory tells us that these responses are driven by 'psychological' intuitions. And this allows us to construct a complementary argument that suggests that utilitarian responses are unlikely to fall victim to a parallel argument because they do not rely on 'psychological' intuitions. This is how "dual-process theory explains why a parallel argument casting doubt on consequentialism is unlikely to go through."[27]

The main steps of Greene's reasoning thus appear to be roughly as follows:

> P1) The deontological responses that have been shown to be sensitive to morally irrelevant factors are based on 'psychological' intuitions.
> P2) Utilitarian responses are not based on 'psychological' but 'philosophical' intuitions.
> C) A parallel argument against utilitarian judgments is therefore unlikely to succeed.

But an argument along these lines, even when we fill in the gaps, is not very convincing. To begin with, the argument presumes that the unreliability of deontological responses is due to their being driven by

[21] See Greene [1], p. 14; [3], pp. 711–713.
[22] Berker [14], p. 326.
[23] Greene [1], p. 18.
[24] See in particular Cushman et al. [22]; Hauser et al. [23]. Note though that the evidence provided by these studies is rather mixed and limited.
[25] Greene [1], p. 20.
[26] Greene [1], p. 20.
[27] Greene [1], p. 15.

'psychological' intuitions. And it is unclear whether this is the case. The fact that the unreliable intuitions are 'psychological' intuitions does not mean that they are unreliable *because* they are 'psychological' intuitions. Their 'psychological' nature could be completely unrelated to their sensitivity to irrelevant factors. The fact that utilitarian judgments differ psychologically from deontological judgments would then be beside the point. Greene does relatively little to explain why the fact that deontological intuitions are sensitive to morally irrelevant factors should be due to their being 'psychological' intuitions. If anything, he appears to suggest that the same considerations that underlie the functionalist argument explain why specifically 'psychological' intuitions are unreliable. However, as will become clear in the discussion of the functionalist argument, this is confused. The functionalist argument and the argument from morally irrelevant factors are on two different levels, and considerations underlying the former cannot inform latter.

But even if we could say that the sensitivity to irrelevant factors of deontological intuitions is due to their being 'psychological' intuitions, this would not allow us to rule out that utilitarian judgments may be vulnerable to a parallel argument. The fact that responses that are based on 'psychological' intuitions tend to be sensitive to irrelevant factors because they are based on 'psychological' intuitions simply does not entail that responses that are not based on 'psychological' intuitions are unlikely to be sensitive to irrelevant factors (or insensitive to relevant factors). For they may of course have this defect *in spite of* not being based on 'psychological' intuitions. And importantly, this is more than just a hypothetical possibility. As noted above, there are at least concrete reasons to suppose that some utilitarian judgments *are* insensitive to morally relevant factors. I am not here positively asserting that such factors as kinship *definitely are* morally relevant. Kumar and Campbell have rightly pointed out that such claims are to some extent controversial. Rather, my point is that these claims are at least *reasonably plausible*. They are too plausible to be brushed aside on the grounds that utilitarian responses differ psychologically from those intuitions that have already been experimentally demonstrated to be responsive to irrelevant factors. When a deontologist rejects some consequentialist judgment as failing to account for a morally relevant factor (e.g. kinship), it simply does not do to respond that this cannot be true because the judgment is not based on a psychological intuition. Appeals to what dual-process theory might predict about whether utilitarian judgments are open to a similar objection are simply way too speculative to be of any dialectical use in this situation.

This means two things: The moral significance of dual-process theory remains unclear, and the problem that utilitarian judgments are plausible targets for a parallel argument persists.

The functionalist argument

More recently, Greene has sketched another way of deriving normative conclusions from dual-process theory. Greene refers to it as the "indirect route"[28]; I call it the functionalist argument. The basic idea is that we are facing two types of moral problems – unfamiliar and familiar ones – that must be dealt with in two different ways that correspond to the two processes identified by dual-process theory. Unfamiliar problems are those with which we have "inadequate evolutionary, cultural, or personal experience."[29] Familiar problems are those with which we have adequate such experience. When facing familiar moral problems, we can rely on our automatic, deontological gut reactions, because they have over time adjusted to these problems through evolutionary, cultural and personal learning processes. By contrast, when we are confronting peculiarly modern, unfamiliar problems, we must distrust our automatic mode and switch into the 'manual' mode, which yields utilitarian solutions.

The moral problems we are facing range from the difficulty of cooperative behavior in everyday life to more complex and/or recent problems such as violent conflict, global warming, terrorists using weapons of mass destruction, global poverty, bioethical problems, the place of religion in public live, capital punishment, abortion, and so forth.[30] Familiar problems are often what Greene calls 'Me vs Us' problems. These are problems associated with conflicts between individuals within the same group. In order to reap the benefits of cooperation, individuals must sometimes restrain their own selfishness, especially in Prisoner's Dilemma-like situations. Morality enables cooperation in these situations by telling people to put

[28] Greene [3]; see also Greene [1, 2, 4]. For an interesting discussion, see Lott [24].
[29] Greene [3], p. 714.
[30] See e.g. Greene [2], pp. 98–99, 348; [4], p. 73.

'Us' ahead of 'Me'. Unfamiliar problems, by contrast, are often of the 'Us vs Them' type. They concern conflicts between groups. 'Us vs Them' problems have two dimensions: "First, there is plain old selfishness at the group level, also known as *tribalism*. Humans nearly always put Us ahead of Them. Second, beyond tribalism, groups have genuine differences in values, disagreements concerning the proper *terms* of cooperation."[31] And both sub-problems entail costly inter-tribal conflicts.[32] Problems arising from recent technological or social developments – e.g. bioethical problems − are also typical examples of unfamiliar problems.

Here is a rough formal summary of the functionalist argument:

P1. We are facing two types of moral problems, familiar and unfamiliar ones.
P2. Our moral cognition operates in two modes, an automatic and a manual mode.
P3. Our automatic mode has evolved to deal with familiar problems but not with unfamiliar ones.
P4. The automatic mode yields deontological judgments and the manual mode yields consequentialist ones.
P5. We can rely on our automatic responses only when facing problems that these responses have evolved to deal with.
C. We can rely on our automatic, deontological responses when dealing with familiar problems but must switch to manual, consequentialist reasoning when dealing with unfamiliar problems.

It is important to note that Greene has, somewhat surreptitiously, shifted the topic of his inquiry. The functionalist argument is meant to determine which moral norms we should adopt in order to achieve certain pre-defined goals. The suggestion to rely on our automatic, deontological responses in familiar but not in unfamiliar situations is best understood as a heuristic.[33]

By contrast, both the argument from evolutionary history and the argument from morally irrelevant factors were attempts to identify the moral goals we should try to achieve in the first place. This explains why the conclusions of these arguments differ. The conclusion of the functionalist argument is rather conciliatory. It is an argument for utilitarianism and against deontology, but only with regard to some moral problems. It even involves a partial vindication of deontology, namely as a solution to familiar moral problems. This contrasts with Greene's earlier evolutionary skepticism about the reliability of our deontological intuitions and with his hope that the argument from morally irrelevant factors will eventually undermine *all* deontological intuitions.[34]

The functionalist argument cannot, I think, be criticized for failing to make use of dual-process theory. But there are other reasons to be skeptical that it constitutes a compelling way of drawing normative conclusions from dual-process theory. Below, I highlight five interrelated difficulties with the functionalist argument and Greene's statement of it.

First, since the functionalist argument is about achieving moral goals, it already presupposes answers to contested moral questions. When we ask what norms best serve a given purpose, we are asking an essentially *instrumental* question.[35] This implies that we must already have an idea of what our final (non-instrumental) goals ought to be. The functionalist argument does therefore not provide an answer to the deontology/utilitarianism controversy, understood as a controversy about what is finally morally valuable. Rather, it *presupposes* an answer to this question. This issue is less serious when the moral problems identified by Greene are generally agreed to really be problems. For instance, deontologists and utilitarians alike can agree that weapons of mass destruction and climate change *are* problems, so there is nothing question begging about maintaining that we must find ways of overcoming *these* problems. But as soon as we turn to more controversial issues – especially those contested between deontologists and utilitarians – the functionalist approach threatens to beg the question.[36]

Second, Greene does not seem to be sufficiently aware of the discontinuity between the functionalist

[31] Greene [2], pp. 66–67, see also Greene [2], p. 99; [4], pp. 72–73. His characterization of 'Us vs Them' problems as 'unfamiliar' is problematic, though (see note 45 below).
[32] Greene [2], pp. 1–27, pp. 293–295.
[33] Bruni et al. call this the 'collective usefulness' view: "According to this view, certain forms of moral thinking are to be recommended because they serve instrumentally to further widely shared goals, such as a reduction in conflict, or an increase in social cohesion." ([25], p. 160). Note that there is a long tradition in utilitarian thought of embracing at least some common-sense moral rules as useful rules-of-thumb (Sunstein [26], p. 533).
[34] Greene [1], p. 21.
[35] See note 33 above.
[36] Greene seems to be aware of this problem and promises to address it in his book ([1], p. 24). But as I explain below, I find his treatment of these issues in his book unconvincing.

argument and his other arguments. This is evidenced by how Greene characterizes the relation between the argument from morally irrelevant factors and the functionalist argument. The argument from morally irrelevant factors is clearly informed by the conventional rather than the functionalist approach to morality. The factor 'personalness' is dismissed as lacking *intrinsic* moral significance rather than as being irrelevant with regard to the achievement of cooperation (or some other moral goal). This means that the genealogical considerations that underlie the functionalist argument do not predict that (or explain why) deontological responses are sensitive to factors that strike us as morally irrelevant. This, however, is what Greene appears to be insinuating (as briefly mentioned in section 3). He seems to think of the 'indirect' route as an expansion or elaboration of the 'direct route' (the argument from morally irrelevant factors). The latter allowed "[l]imited progress", whereas the former offers "a more general theory that tells us when our judgments are likely to go astray."[37] But this is confused. The genealogical considerations that inform the functionalist argument predict, for instance, that our intuitive responses foster within-group cooperation while hindering between-group cooperation. Our conception of moral relevance, however, is distinct from the question of what is instrumentally necessary for the achievement of cooperation. Therefore, the genealogical considerations underlying the functionalist argument imply nothing about our intuitions' sensitivity to morally irrelevant factors.

It might be possible to disentangle these two arguments and to assign each a meaningful role in the overall structure of the argument. But Greene's own understanding and presentation of how the functionalist argument relates to the rest of his empirically informed case for utilitarianism seems confused. This confusion is also apparent in his discussion of the incest thought experiment, one of his favored illustrations of how science can advance moral philosophy.[38] It involves a brother and a sister who engage in a consensual romantic relationship using adequate birth control. Greene's argument is based on the plausible causal premise that people tend to condemn incest because incest led to genetic defects in the environment in which our ancestors evolved. He then maintains that "[w]hether or not a behavior increased the probability of deleterious consequences in the environment of our ancestors is irrelevant to its present moral acceptability, so long as this behavior does not also causes [*sic*] similar harm in our present environment" (which includes methods of birth control).[39] This leads him to conclude that we ought not condemn their behavior. While this may be true, Greene presents this as an illustration of the argument from morally irrelevant factors. He writes that the causal premise "tells us that people's judgments are, in this instance, determined by their sensitivity to a morally irrelevant factor."[40] But the causal premise does not specify the factors that trigger the deontological response. Our intuitive aversion to incestuous relationships is not sensitive to this evolutionary fact in the same manner in which our moral judgments are sensitive to personal force. Rather, what is triggering it is probably simply the fact that it is siblings rather than unrelated people who are having a love affair. The incest example is thus either an instance of the functionalist argument (as Greene appears to suggest elsewhere[41]) or of the argument from evolutionary history.

Third, his suggestion that our automatic, deontological responses are reliable in 'familiar' situations is questionable. Greene reasons that "[a]utomatic settings can function well only when they have been shaped by *trial-and-error experience*", be it a biological, cultural or personal one.[42] But this reasoning rests on the problematic assumption that the criterion of selection of these mechanisms is whether a response helps us solve our moral problems.[43] This is roughly true of biological evolution, although only indirectly: Biological evolution does not select for moral dispositions that solve our moral problems but for dispositions that enhance our inclusive fitness. But since cooperative traits increase inclusive fitness and since lack of cooperation is one of the to-be-solved problems, our naturally evolved moral dispositions may be expected to provide fairly

[37] Greene [3], p. 713. Elsewhere, he writes that "whether a judgment is produced by a process that is emotional, heuristic, or a by-product of our evolutionary history is not unrelated to whether that judgment reflects a sensitivity to factors that are morally irrelevant." ([1], p. 12)

[38] The example is from Haidt et al. [27] and Haidt [28].

[39] Greene [1], p. 11.

[40] Greene [1], p. 11. Shortly after, he explicitly endorses the 'argument from morally irrelevant factors'-interpretation suggested by Berker. And he writes that his characterization of the argument from morally irrelevant factors is modelled on the incest argument ([1], p. 15). Elsewhere ([3], p. 712), his presentation of the incest case is also embedded in a discussion of the argument from morally irrelevant factors.

[41] Greene [1], p. 22.

[42] Greene [3], p. 714.

[43] This is also noted by Greene ([3], p. 714)

good solutions to some of our (familiar) moral problems. The case of cultural evolution, by contrast, is much trickier. It is fair to say that the laws of cultural transmission are still poorly understood. But according to one recent suggestion, informed by the epidemiological approach, the 'cultural fitness' of a moral norm is a function of the following three factors:

> "(i) It yields material benefit to its believers or to the members of a culture who are in a position to indoctrinate others;
> (ii) it is situated in a narrative context that is easy to learn because, e.g., it integrates with existing beliefs about the nature of the world or captures the imagination; or
> (iii) it has emotional appeal, due to the intrinsic content of the belief or accompanying practices, such as emotional conditioning or emotionally intense religious rituals."[44]

Of these three factors, only the first factor might plausibly be connected to the resolution of moral problems. But even this is doubtful. For one thing, if some norms exist because they benefit the powerful indoctrinators, they are probably poorly aligned with the correct moral values. They do not solve our moral problems but the prudential problems of the powerful indoctrinators, so to speak. For another, if a norm benefits only the believers, this is bound to happen at the expense of members of the outgroup. It is, for instance, beneficial to believe that it is morally permissible to kill, enslave or, indeed, eat members of the outgroup.[45] Finally, to describe our personal learning experience as an adaptive 'trial-and-error process' strikes me as misleading, too. For when we make a bad moral judgment (say, that the above incestuous relationship is a moral abomination), we do not get an 'error message' that prompts us to adjust our judgment. Greene likens the personal moral learning experience to learning to fear hot stoves by touching them.[46] But there is no obvious equivalent to the sensation of heat when we make a wrong moral judgment. You can go your entire life and never notice that you were wrong about abortion, incest, capital punishment, and so forth. In fact, a whole lot of people *do* go their entire lives without noticing that they have been wrong about these issues (whatever the truth about these issues may be).

All this is not to deny that there is some truth in Greene's idea that our automatic responses to familiar problems may contain acquired moral wisdom.[47] But the extent to which they do is unclear. The three learning mechanisms are much less linear and reliable than Greene's talk of a trial-and-error process suggests. The heuristic to use the automatic mode in familiar situations will probably misfire.[48]

Fourth, his positive case for utilitarianism strikes me as confused. Even if we agree that our deontological responses fail us in unfamiliar situations because they have not evolved to deal with these situations, this still does not mean that manual mode reasoning yields good solutions to these problems. And this is because when we engage in manual moral reasoning, we typically do not engage in the kind of functionalist or instrumentalist reasoning that seeks to find 'solutions' to predefined 'problems'. When we contemplate a moral question ("Should I shove the heavy person from the bridge?", "Should we take money from the rich and give it to the poor?", etc.), we are not asking ourselves which steps must be taken to achieve some pre-defined moral goal, such as cooperation. Rather, moral deliberation centrally involves figuring out what is morally valuable in the first place, that is, what our moral goals should be and what moral side constraints we might have to respect. Manual moral reasoning is thus typically concerned with answering a different question than the one that is central to Greene's functionalist argument. Even if manual mode reasoning should support utilitarianism, this would not show that utilitarianism provides workable 'solutions' to our moral 'problems'. If anything, it would show that the maximization of welfare is the only thing that matters morally. But this tells us little about

[44] Prinz [29], p. 220. The epidemiological approach was pioneered by Sperber [30].

[45] See Prinz [29], pp. 223–229. Greene could object that this is an *inter*-tribal conflict and as such not suited for our automatic mode, anyway. Indeed, at one point, Greene writes: "Of course, Us versus Them is a very old problem. But historically it's been a tactical problem rather than a moral one." ([2], p. 15) But this statement is puzzling. What does it mean for a problem to be a tactical rather than moral one? And why is the *intra*-tribal tragedy of the commons (presumably as 'tactical' a problem as one can imagine) a moral problem rather than a tactical one? And does this mean that 'familiarity' is *not* the decisive criterion, at least not the only one? In any case, even if we abstract from this specific problem with the cultural trial-and-error process, other problems remain.

[46] Greene [3], p. 714.

[47] See Railton [31] and Sauer [32].

[48] Similarly, Bruni, Mameli and Rini conclude that Greene's suggested heuristic is unconvincing until it has been corroborated by "a very ambitious empirical research program" ([25], p. 171).

how exactly to achieve, say, inter-tribal cooperation. In order to come up with solutions to unfamiliar problems, we would instead have to switch into the 'social scientist mode', so to speak, which engages in means-end reasoning and which might tell us how to best achieve inter-tribal cooperation and other goals.

The fact (if it is a fact) that 'manual' moral reasoning supports utilitarianism does therefore not constitute a functionalist vindication of utilitarianism. In his *Moral Tribes*, Greene offers two other rationales for choosing utilitarianism as the solution to unfamiliar moral problems.

One is that we should take a pragmatic approach and look out for what he calls a metamorality. This is a moral system that allows adjudicating inter-tribal conflicts because it is based on shared moral values: "This is the essence of deep pragmatism: to seek common ground not where we think it ought to be, but where it actually is."[49] And Greene believes that utilitarianism is particularly well suited to serve as such a metamorality.

The other is that utilitarianism is supported by rational, empirically informed moral theorizing. While he does not directly argue for the *truth* of utilitarianism, he contends that utilitarianism becomes "uniquely attractive once our moral thinking has been *objectively improved* by a scientific understanding of morality."[50] And by this he means that utilitarianism is supported by the argument from morally irrelevant factors[51] and evolutionary debunking arguments,[52] alongside a range of other considerations.

It is easy to be confused by Greene's argumentation. One problem is that the above two criteria are in conflict with each other. If we 'objectively improve' our moral thinking in the way Greene envisages, we are bound to move away from the 'common ground where it actually is'. After all, 'objective improvement' involves, among other things, debunking generally shared deontological intuitions, which is the opposite of starting from common ground where it actually is. Relatedly, it is rather implausible to claim that utilitarianism, which has notoriously counterintuitive implications, rests on an overlapping moral consensus.[53] Thus, either we solve unfamiliar problems by appeal to commonly shared values, in which case utilitarianism will hardly be our morality of choice. Or we solve these problems by relying on 'objectively improved moral thinking'. This method might vindicate utilitarianism, but it will not yield a metamorality based on shared values.

What is more, the rationale behind objectively improving our moral thinking is, in this context, itself rather puzzling. If we are interested in whether norms are conducive to solving pre-defined moral problems, the notion of objectively improving moral thinking as Greene conceives it is difficult to make sense of. Although Greene is reluctant to call this objectively improved morality (utilitarianism) *true*, his argument that empirically informed moral theorizing vindicates utilitarianism proceeds *as though* he was arguing for its truth. For instance, the argument from morally irrelevant factors purports to show that deontological intuitions are not tracking morally significant properties. Other of his arguments are supposed to dispel the impression that utilitarianism has counterintuitive implications.[54] But these considerations are meaningless or irrelevant from the functionalist vantage point. What matters from the functionalist vantage point is whether a given system of moral norms is *functional*, that is to say, whether it helps us overcome our moral problems. It may be true that objectively improved moral thinking supports utilitarianism. But given the functionalist framework, it is unintelligible why we should care about which morality is supported by objectively improved moral thinking in the first place.[55] This undermines Greene's empirically informed case for utilitarianism as the solution to our most pressing moral problems.

Fifth, and lastly, Greene's suggested solution – utilitarianism – is to some extent arbitrary and in fact not even that much of a solution. It is to some extent arbitrary because there are other norms or instructions that can serve the same function just as well as utilitarianism. Take 'Us vs. Them' problems, that is, inter-tribal conflicts arising from selfishness at the group level and differences in values. If this is the problem, we might as

[49] Greene [2], p. 291.
[50] Greene [2], p. 189.
[51] Greene [2], pp. 213–217, p. 261.
[52] Greene [2], pp. 224–245. Note that the evolutionary debunking arguments that feature in Greene's book differ from Greene's earlier evolutionary debunking arguments.
[53] Similarly, Wielenberg [33], p. 914.
[54] Greene [2], pp. 254–285.
[55] Similarly, Tobia [34], p. 749. By contrast, the idea behind seeking shared ground *is* intelligible. Given that one of the to-be-solved problems are the conflicts resulting from disagreement, identifying shared values may be a way of mitigating these conflicts. Interestingly, however, Greene himself appears to favor an epistemological rationale, which is less intelligible given the functionalist framework ([2], pp. 188–189).

well simply establish moral norms that tell people 1) to avoid in-group-favoritism and be generous towards the outgroup, and 2) to be respectful or tolerant of the views and practices of other 'tribes'. These two norms would be the most obvious and straightforward moral solutions to these problems. Similar *ad hoc* solutions can easily be formulated for other problems that Greene thinks need to be solved. There seems to be no need to accept a grand moral theory such as utilitarianism. And this points to the other problem: it is doubtful whether utilitarianism, or any *ad hoc* norm of the above sort, really qualifies as a 'solution'. For the problem is not so much that we do not know which norms, if complied with, would solve our 'moral problems'. As just seen, it should not be too difficult to formulate norms that fit this description. Rather, the main problem is the compliance itself, that is, getting people to actually do the things that need to be done in order to overcome the problems. It is instructive here to compare Greene's proposal to that of Ingmar Persson and Julian Savulescu, who start from a very similar diagnosis. They, too, are concerned that commonsense morality, evolved as a solution to the cooperation problems of relatively small groups, has become dysfunctional as a result of recent technological developments. Persson and Savulescu are particularly concerned about weapons of mass destruction and climate change. Their suggested solution, however, is the biological enhancement of people's moral *motivation*. And this makes more sense. Although compliance with some moral system as, say, utilitarianism would certainly prevent the use of weapons of mass destruction, the exhortation to accept utilitarianism (and, by implication, to refrain from mass murder) would do very little to avert the use of such weapons. It will simply not be heeded. Likewise, the proposal to settle inter-tribal ideological conflicts by converging on utilitarianism as a shared morality is just too unlikely to gain sufficient traction among the members of the various ideological camps to actually qualify as a 'solution'. Few policy makers will be impressed with the suggestion to resolve conflicts between, say, Christians and Muslims by encouraging them to jointly embrace utilitarianism. At least with regard to some of the problems in question, Greene's solution is arguably a case of what David Estlund has called hopelessly aspirational theory. Hopelessly aspirational theory gives normative instructions that *could* be complied with but that we know *will not* be complied with. While hopelessly aspirational theory may be philosophically legitimate, as Estlund believes, it is just not the kind of solution we are looking for when we actually want to solve real-life problems.[56]

In summary, Greene's functionalist argument raises more questions than it answers. It makes use of dual-process theory, but it is plagued by internal tensions and incongruities. The functionalist argument, at least in its current form, does not constitute a compelling way of extracting normative conclusions from dual-process theory. Whether an improved version of the argument might allow deriving useful moral heuristics from dual-process theory remains to be seen.

The argument from confabulation

One final option for Greene would be to embrace the argument from morally irrelevant factors, but to insist that the neuroscientific findings inform not this argument but the argument from confabulation, which complements the argument from morally irrelevant factors. To be sure, we noted above that the argument from morally irrelevant factors might be deployed against utilitarian judgments, too. But this does not mean that he cannot use it to debunk *some* deontological intuitions and then dismiss theories that confirm these intuitions as confabulatory post hoc rationalization. This would allow Greene to accept Berker's observation that neuroscience plays no role in the argument from morally irrelevant factors and yet refute the charge that neuroscience is normatively insignificant.[57] Let me explain:

The argument from morally irrelevant factors (just like the argument from evolutionary history and parts of the functionalist argument) is meant to show that deontological intuitions are unreliable. But Greene is aware that not all deontological theories appeal to these intuitions. Immanuel Kant's moral theory is just one of many exceptions. These independent, often fairly sophisticated justifications of deontology appear to be unaffected

[56] Persson and Savulescu [35]; Estlund [36].

[57] Surprisingly, neither Berker nor Greene consider this option, even though Greene's presentation of the argument clearly invokes the neuroscientific findings. Berker only remarks that there is no need to discuss this argument, as it presupposes the success of the debunking of deontological intuitions, which Berker disputes ([14], 315). But this is rather uncharitable. As just noted, Greene can still use the argument from morally irrelevant factors to debunk *some* deontological intuitions (while admitting its limitations) and then combine it with the argument from confabulation. If the latter uses the neuroscientific findings, this would suffice to refute Berker's main criticism.

by the debunking attack on deontological gut reactions. Therefore, in a second step, Greene contends that these more sophisticated justifications of deontology are merely post hoc rationalizations of our intuitive gut reactions. To illustrate the logic of post hoc rationalization, Greene presents the story of Alice. According to Alice's own account, her evaluation of potential romantic partners is based on such attributes as their intelligence, their sense of humor, their likability, and so forth. Strangely enough, though, her judgments happen to be perfectly predicted by the men's heights. Given the well-documented human tendency to engage in confabulatory post hoc rationalization, the most plausible interpretation of these 'data' is that her judgments are really determined by a preference for tall men while her own explanations of her judgments are merely confabulatory post hoc rationalizations.[58] Greene believes that deontological philosophers are like Alice. They offer all sorts of elaborate justifications of deontology, but they are really just engaged in rationalizing their gut reactions. This is why the fact that there are deontological theories that do not just appeal to the debunked deontological intuitions provides little comfort for deontologists.[59]

Now, Greene could retort that the neuroscientific findings are normatively significant because they play a role in this argument from confabulation, which complements the debunking of our deontological intuitions. The neuroscientific findings reveal the "factor, namely emotional response, that predicts deontological judgment."[60] According to Greene, just as height predicts Alice's judgment, so the level of emotional arousal predicts the claims of deontological theory. Greene observes that there are, for instance, complicated deontological theories that explain why it is appropriate in SWITCH to sacrifice one person to save five others but not in FOOTBRIDGE, and this claim just so happens to be predicted by whether these scenarios elicit a strong emotional reaction or not.[61] It is this suspicious coincidence that makes the post hoc rationalization thesis so plausible. And the neuroscientific research played the decisive role in uncovering the link between emotional arousal and deontological judgment.[62]

The problem with this response, however, is that using emotional arousal as the predictor is to make an unnecessary detour. Greene is claiming that the judgments that deontological theorists are rationalizing reflect their moral *intuitions*: "their reasoning serves primarily to justify and organize their preexisting *intuitive conclusions* about what's right or wrong."[63] But this means that these intuitions are an even more convenient predictor of the claims of deontological theory: The best support for the post hoc rationalization thesis is provided by the suspicious coincidence that deontological theorizing tends to confirm our intuitions, such as that we should hit the switch in SWITCH but spare the heavy person in FOOTBRIDGE. It may be true that the level of emotional arousal that a moral question provokes predicts the claims of deontological theory, too. But it does so because it determines our intuitive responses, which in turn underlie the judgements that deontologists are engaged in rationalizing. This makes it difficult to see why we should not use our intuitive responses as the predicting factor straightaway. There is simply no point in going one step further back and identifying the factor that predicts our intuitive judgments. It would be as if we did not use the men's heights as the predicting factor of Alice's judgments but instead, say, the set of alleles that determine their heights. This would make genetics play a role in the argument, but only by making things needlessly complicated. Similarly, using emotional arousal as the predictor only makes things needlessly complicated. It is easier to check directly whether the claims of deontological theory are predicted by people's intuitive responses − whatever the level of emotional arousal.

It seems, then, that the normative significance of the neuroscientific research cannot be established by stressing its role in the argument from confabulation, either. Might the argument from confabulation be sound nonetheless, even though it does not rely on the neuroscientific results? I am inclined to acknowledge that there is some truth in the suspicion that deontological theory is post hoc rationalization, but the precise extent to which this is the case is difficult to estimate.[64]

[58] On post hoc rationalization, see e.g. Uhlmann et al. [37]; Wheatley and Haidt [38]; Wilson [39].

[59] Greene [10], pp. 60–63, 67–72; [3], p, 718; for instructive discussions, refer to Dean [40], pp. 47–48; Mihailov [41].

[60] Greene [10], p. 68.

[61] Greene [10], p. 68.

[62] In particular Greene et al. [5].

[63] Greene [3], p. 718, emphasis added. Note that deontologists are not claimed to only rationalize their deontological intuitions but *both* their deontological as well as their consequentialist intuitions. After all, deontologists provide explanations of why FOOTBRIDGE is morally different from SWITCH, rather than just why FOOTBRIDGE calls for a deontological response.

[64] Besides, the argument is an ad hominem attack, which, even if sound, has no place in scholarly debate [42].

Conclusion and some remarks on the normative significance of other empirical findings

It must be concluded that either the neuroscientific findings are not needed for Greene's argument or that the argument, while reliant on these findings, is unconvincing: The argument from morally irrelevant factors suffers from a mixture of these two defects. The argument itself does not rely on dual-process theory, as correctly observed by Berker. It also suffers from the defect that it might be deployed against utilitarian judgments, too. Greene has offered a response that could potentially remedy both problems: He has maintained that dual-process theory predicts that such a parallel argument against utilitarianism is unlikely to go through. But we have found that this complementary argument is unpersuasive. The functionalist argument, by contrast, relies squarely on dual-process theory, but it is riddled with other problems. I also looked at the argument from confabulation. It appears to be based on Greene and colleagues' neuroimaging results, but this appearance proved deceptive.

Dual-process theory possesses great intrinsic significance as a psychological theory, and Greene and his colleagues deserve praise for their pioneering work. But as of now, its normative significance remains unclear. While we should not rule out in principle that dual-process theory and Greene's neuroscientific findings can advance normative theorizing, he has so far failed to show how. Interestingly, Greene seems at times prepared to shift the focus from neuroscience and dual-process theory to other findings from experimental moral psychology. To be sure, he is adamant that the "neuroscientific data have implications for normative ethics" and that "the dual-process theory of moral judgment *is* essential to [his] argument".[65] But he also places great emphasis on the fact that the argument from morally irrelevant factors relies centrally on his and his colleagues' study on which factors of a sacrificial dilemma make it trigger deontological responses. At one point, he even asserts that this study is "the one with the most direct relevance to normative issues".[66] So even if the neuroscience and dual-process theory should be normatively insignificant, this would not show that findings from experimental moral psychology are doing no work in the argument from morally irrelevant factors at all. Greene believes that this takes some of the sting out of Berker's criticism.[67]

The question of the present paper was whether dual-process theory and the supporting neuroscientific findings have any moral implications. I will not here offer a comprehensive discussion of whether such survey findings can advance normative theorizing in the way imagined by Greene, which is a separate issue.[68] But I do wish to note, by way of conclusion, that there is something startling about Greene's apparent readiness to shift the focus from his neuroscientific research and dual-process theory to the survey findings, precisely because these are two separate issues. Berker's objection concerned in the first instance the normative insignificance of neuroscience and dual-process theory, rather than that of findings from experimental moral psychology in general.[69] Therefore, retorting that Berker failed to appreciate the normative significance of the survey findings is somewhat beside the point.[70] And whether specifically the neuroscientific findings and dual-process theory are normatively significant appears to be the dialectically important question. The dual-process model is Greene's chief contribution to our understanding of the psychology of moral judgment. And he has been eager from the outset to derive normative conclusions from this line of research.[71] It would be distinctly underwhelming if what is really normatively significant are, in the end, the somewhat less spectacular survey

[65] Greene [1], pp. 8 and 14. His 2014 paper is also obviously intended to demonstrate the moral significance of dual-process theory.

[66] Greene [1], p. 4.

[67] Greene [1], p. 17.

[68] For two interesting discussions, see Kahane [43] and Kumar and Campbell [21], pp. 315–319. Rini [44], too, defends the method underlying the argument from morally irrelevant factors.

[69] Hence the title of his paper, "The Normative Insignificance of Neuroscience". See also in particular Berker [14], pp. 294, 325–327.

[70] And Berker can hardly be faulted for not appreciating the normative significance of this sort of experimental moral psychology given that Greene did not much emphasize its normative significance until after Berker had raised concerns about the normative significance of neuroscience. The focus had no doubt been on the neuroscientific findings and dual-process theory. The question which factors trigger deontological responses initially played only a subordinate role. It was necessary to make a provisional guess on this question in order to test the dual-process hypothesis. ([5], p. 2107; see also [1], p. 27; [6]; [3], p. 701 n17). It was only later that Greene and colleagues set out to develop a more precise account of the principles that govern people's responses to trolley dilemmas [45]. This study had already been published when Berker wrote his article, and he mentions it in a footnote ([14], p. 323 n73). But Greene's most complete statement of why the empirical findings matter, his 'The Secret Joke of Kant's Soul' [10], predates this study and does not take it into account.

[71] Greene [46], p. 849. Greene initially took the findings to refute moral realism rather than deontology.

data, which are independent from dual-process theory. As Greene rightly points out, the question of exactly which factors trigger deontological responses and dual-process theory are two different things: "the personal/impersonal distinction tentatively posited in 2001 and the dual-process theory of moral judgment are completely orthogonal ideas. The dual-process theory could be completely correct, even if the personal/impersonal distinction is completely wrong, and vice versa."[72] This also applies to the updated 2009 theory about which factors deontological responses are sensitive to. To shift the philosophical focus to these survey data is thus not to shift it to a different aspect of dual-process theory. It is to shift it to a line of experimental research that is largely independent from dual-process theory and thus from Greene's primary and most innovative contribution to moral psychology. Even if it should be true that such survey findings may advance moral theory, it would be quite anticlimactic if this were the only way in which Greene's empirical work is normatively relevant.

Acknowledgements I would like to thank Katharina Brecht, Sabine Döring, Malte Hendrickx, Michael Wenzler and the anonymous referees of this article.

Funding Studienstiftung des deutschen Volkes.

References

1. Greene, J. 2010. *Notes on 'The Normative Insignificance of Neuroscience' by Selim Berker*. Unpublished manuscript.
2. Greene, J. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York: Penguin Press.
3. Greene, J. (2014). Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. *Ethics* 124 (4), 695–726.
4. Greene, J. 2017. The rat-a-gorical imperative: Moral intuition and the limits of affective learning. *Cognition* 167 (1): 66–77.
5. Greene, J., B.R. Sommerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen. 2001. An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science* 293 (5537): 2105–2108.
6. Greene, J. 2009. Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Psychology* 45 (3): 581–584.
7. McGuire, J., R. Langdon, M. Coltheart, and C. Mackenzie. 2009. A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology* 45 (3): 577–580.
8. Greene, J., 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 (2): 389–400.
9. Klein, C. 2011. The Dual Track Theory of Moral Decision-Making: a Critique of the Neuroimaging Evidence. *Neuroethics* 4 (2): 143–162.
10. Greene, J. 2008. The Secret Joke of Kant's Soul. In *Moral Psychology: Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. W. Sinnott-Armstrong, 35–80. Cambridge, MA: MIT Press.
11. Greene, J. 2005. Cognitive Neuroscience and the Structure of the Moral Mind. In *Innateness and the Structure of the Mind: Volume 1*, ed. P. Carruthers, S. Laurence, and S. Stich, 338–352. New York: Oxford University Press.
12. Greene, J. 2005. Emotion and Cognition in Moral Judgment: Evidence from Neuroimaging. In *Neurobiology of Human Values*, ed. J.-P. Changeux, A. Damasio, W. Singer, and Y. Christen, 57–66. Berlin/Heidelberg: Springer.
13. Kahane, G. 2011. Evolutionary Debunking Arguments. *Noûs* 45 (1): 103–125.
14. Berker, S. 2009. The Normative Insignificance of Neuroscience. *Philosophy and Public Affairs* 37 (4): 293–329.
15. Sauer, H. 2012. Morally irrelevant factors: What's left of the dual-process model of moral cognition. *Philosophical Psychology* 25 (6): 783–811.
16. Singer, P. 2005. Ethics and Intuitions. *The Journal of Ethics* 9 (3–4): 331–352.
17. Kahane, G. 2014. Evolution and Impartiality. *Ethics* 124 (2): 327–341.
18. Mason, K. 2011. Moral Psychology and Moral Intuitions: A Pox On All Your Houses. *Australasian Journal of Philosophy* 89 (3): 441–458.
19. Tersman, F. 2008. The reliability of moral intuitions: A challenge from neuroscience. *Australasian Journal of Philosophy* 86 (3): 389–405.
20. Greene, J. 2016. Solving the Trolley Problem. In *A Companion to Experimental Philosophy*, ed. J. Sytsma and W. Buckwalter, 175–189. Malden, MA: Wiley Blackwell.
21. Kumar, V., and R. Campbell. 2012. On the normative significance of experimental moral psychology. *Philosophical Psychology* 25 (3): 311–330.
22. Cushman, F., L. Young, and M. Hauser. 2006. The Role of Conscious Reasoning and Intuition in Moral Judgment: Testing Three Principles of Harm. *Psychological Science* 17 (12): 1082–1089.
23. Hauser, M., Cushman, F., Young, L., Kang-Xing, J., & Mikhail, J. (2007). A Dissociation Between Moral Judgments and Justifications. *Mind & Language*, 22(1), 1–21.
24. Lott, M. 2016. Moral Implications from Cognitive (Neuro)Science? No Clear Route. *Ethics* 127 (1): 241–256.
25. Bruni, T., M. Mameli, and R.A. Rini. 2014. The Science of Morality and its Normative Implications. *Neuroethics* 7 (2): 159–172.
26. Sunstein, C. 2005. Moral heuristics. *Behavioral and Brain Sciences* 28 (4): 531–542.

[72] Greene [1], p. 27; see also Greene [6].

27. Haidt, J., F. Bjorklund, and S. Murphy. 2000. *Moral dumbfounding: When intuition finds no reason*. Unpublished manuscript.
28. Haidt, J. 2001. The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review* 108 (4): 814–834.
29. Prinz, J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.
30. Sperber, D. 1996. *Explaining Culture. A Naturalistic Approach*. Oxford: Blackwell.
31. Railton, P. 2014. The Affective Dog and Its Rational Tail: Intuition and Attunement. *Ethics* 124 (4): 813–859.
32. Sauer, H. 2012. Educated intuitions. Automaticity and rationality in moral judgment. *Philosophical Explorations* 15 (3): 255–275.
33. Wielenberg, E.J. 2014. Joshua Greene: Moral Tribes: Emotion, Reason, and the Gap between Us and Them. *Ethics* 124 (4): 910–916.
34. Tobia, K. 2015. Moral Tribes: Emotion, Reason, and the Gap between Us and Them. *Philosophical Psychology* 28 (5): 746–750.
35. Persson, I., and J. Savulescu. 2012. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press.
36. Estlund, D. (2014). Utopophobia. *Philosophy & Public Affairs,* 42(2), 113–134.
37. Uhlmann, E.L., D.A. Pizarro, D. Tannenbaum, and P.H. Ditto. 2009. The motivated use of moral principles. *Judgment and Decision making* 4 (6): 476–491.
38. Wheatley, T., and J. Haidt. 2005. Hypnotic Disgust Makes Moral Judgments More Severe. *Psychological Science* 16 (10): 780–784.
39. Wilson, T.D. 2002. *Strangers to ourselves: discovering the adaptive unconscious*. Cambridge, MA/London: Harvard University Press.
40. Dean, R. 2010. Does Neuroscience Undermine Deontological Theory? *Neuroethics* 3 (1): 43–60.
41. Mihailov, E. 2015. Is Deontology a Moral Confabulation. *Neuroethics* 9 (1): 1–13.
42. Königs, P. 2018. Two types of debunking arguments. *Philosophical Psychology* 31 (3): 383–402.
43. Kahane, G. 2013. The armchair and the trolley: an argument for experimental ethics. *Philosophical Studies* 162 (2): 421–445.
44. Rini, R.A. 2013. Making Psychology Normatively Significant. *The Journal of Ethics* 17 (3): 257–274.
45. Greene, J., 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 (3): 364–371.
46. Greene, J. 2003. From neural 'is' to moral 'ought': what are the moral implications of neuroscientific moral psychology? *Nature Reviews Neuroscience* 4 (10): 847–850.