

Moral Implications from Cognitive (Neuro)Science? No Clear Route*

Micah Lott

Joshua Greene argues that cognitive (neuro)science matters for ethics in two ways, the “direct route” and the “indirect route.” Greene illustrates the direct route with a debunking explanation of the inclination to condemn all incest. The indirect route is an updated version of Greene’s argument that dual-process moral psychology gives support for consequentialism over deontology. I consider each of Greene’s arguments, and I argue that neither succeeds. If there is a route from cognitive (neuro)science to ethics, Greene has not found it.

As Joshua Greene says, “The cognitive science of ethics is booming.”¹ This raises a question: what significance, if any, does empirical research concerning our moral judgments have for normative ethics? Greene has recently argued that cognitive science can have moral implications in two different ways, which he labels “the direct route” and “the indirect route” (711, 713). Moreover, Greene argues that the indirect route favors certain forms of consequentialism over nonconsequentialist moral theories.

I will argue that Greene is wrong about both the direct and indirect routes. Neither of Greene’s arguments shows that cognitive science has

* For helpful comments on earlier versions of this essay, I thank two anonymous reviewers for *Ethics*, and also Anne Baril, Keith Cox, Daniel Groll, Richard Kim, Jennifer Lockhart, Daniel McKaughan, Liane Young, and the members of my class, What Is Moral Knowledge?, at Boston College.

1. Joshua D. Greene, “Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics,” *Ethics* 124 (2014): 695–726, 695. Parenthetical references in the text are to this article.

normative implications in the ways that Greene claims, including support for consequentialism. If there is a route from cognitive science to normative ethics, Greene has not found it.

Greene's indirect route is a revised version of his earlier empirically based argument against deontology.² Greene's earlier argument was subject to a variety of criticisms.³ In his revised argument, Greene attempts to deal with those criticisms, and what he now labels "the indirect route" differs from his earlier argument. For this reason, I will focus on Greene's revised view. However, I will also note a key change in Greene's argument where that matters for my critique.

I. THE COGNITIVE SCIENCE OF ETHICS: DUAL-PROCESS THEORY

Greene holds that humans have a "dual-process brain," analogous to a camera with both automatic and manual modes. On the one hand, humans have "a variety of automatic settings—reflexes and intuitions that guide our behavior, many of which are emotional" (696). In general, we are not conscious of the processes that trigger these responses, though we usually rely on automatic settings and they are usually good guides. On the other hand, humans have "a manual mode" that is "a general-purpose reasoning system, specialized for enabling behaviors that serve long(er) term goals, that is, goals that are not automatically activated by current or environmental stimuli or endogenous somatic states" (696–97). In contrast to automatic settings, manual mode processes are typically conscious, voluntary, and effortful.

This structure applies to moral judgment: "Moral judgment is influenced both by automatic emotional responses (automatic settings) and controlled, conscious reasoning (manual mode). Moreover, these processes are enabled by the usual cast of neural characters in their characteristic roles" (698).

More details of dual-process theory are important to Greene's indirect route, and I return to them below in Section III. First, however, I will consider Greene's direct route.

2. Joshua D. Greene, "The Secret Joke of Kant's Soul," in *Moral Psychology*, vol. 3, *The Neuroscience of Morality: Emotion, Disease, and Development*, ed. Walter Sinnott-Armstrong (Cambridge, MA: MIT Press, 2008), 35–79.

3. For criticisms of Greene's earlier work, see: Selim Berker, "The Normative Insignificance of Neuroscience," *Philosophy and Public Affairs* 37 (2009): 293–329; F. M. Kamm, "Neuroscience and Moral Reasoning: A Note on Recent Research," *Philosophy and Public Affairs* 37 (2009): 330–45; Guy Kahane and Nicholas Shackel, "Methodological Issues in the Neuroscience of Moral Judgment," *Mind and Language* 25 (2010): 561–82; Richard Dean, "Does Neuroscience Undermine Deontological Theory?" *Neuroethics* 3 (2010): 43–60.

II. IS THERE A DIRECT ROUTE?

The overall question is: does cognitive science have any moral implications? Greene's answer is "yes." But he recognizes that many are skeptical, and he aims to demonstrate these implications without any "is/ought sleight of hand" (711).

At the beginning of his section on the direct route, Greene highlights one way that empirical science can contribute to important normative conclusions. We begin with a normative question about whether people are making good judgments. Scientific research then provides information about how people, in fact, make decisions—that is, what factors their judgments are actually sensitive to. We can combine this information with a normative assumption about what factors they *ought* to be sensitive to, arriving at a conclusion about how good or bad their actual decision making is. Greene gives the following example:

- 1) Capital juries are sensitive to race (scientific information).
- 2) Capital juries ought not be sensitive to race (normative assumption).
- 3) Thus, capital juries, at least sometimes, make bad decisions.

I do not dispute that empirical science can be relevant to our normative judgments in this way. But as Greene quickly acknowledges, someone might insist that "the science, while not insignificant, is *normatively* insignificant. The science does not challenge anyone's values. Instead, it simply alerts us to an application of the values we already have" (712).

Greene aspires to show that science can matter for ethics in a deeper way than this. Thus he offers another example, and this example appears to illustrate better what Greene means by "the direct route," because in this example the science is employed to provide a direct challenge to values. Again we begin with a normative question, this time about consensual adult incest: ought we condemn all incestuous behavior? Science tells us something about the nature and origin of our inclination to condemn. In particular, science says that "the inclination to condemn incest of all kinds is based on an emotional response whose function is to avoid producing offspring with genetic diseases" (712). We then encounter a situation of incest with no risk of genetic diseases. This is a case like Joe and Jane, who are separated in childhood, later meet and fall in love, and who take precautions to protect against producing offspring with genetic diseases. We might be inclined to condemn their behavior. But we can form a normative assumption about whether to rely upon, or resist, our inclination to condemn. Greene suggests the following assumption: "If our inclination to condemn Joe and Jane's behavior depends on an emotional response that makes their behavior seem wrong, and this emotional

response evolved to prevent birth defects, and birth defects are not a special issue in their case, and we have no additional reason to condemn their behavior, then we ought not to condemn their behavior" (712). Combining this normative assumption with what science tells us about our inclination to condemn incest, we can conclude that we should reject our inclination and should not condemn them. This example, Greene says, shows how science "genuinely challenges some people's moral values" (712).

I believe that Greene's treatment of this case is confused, and that the scientific account of our inclinations has no essential role to play here, either in challenging or affirming moral values. To see why, return to the original question: ought we condemn all incestuous behavior? In asking this question, we might find that we are inclined to condemn. We might recoil from all incest. It might strike us as bad. That is a fact about our inclinations. In asking the question, "ought we condemn all incest?" we are asking whether to take those inclinations seriously—that is, whether to adopt the position that they are pushing us toward. This is asking for a justification for condemning all incestuous behavior.

To answer the question about justification, we need some account of bad-making or good-making features of incest that would serve as grounds for condemning, or not condemning, all (or any) cases. Is incest inconsistent with proper respect for persons? Does it erode important psychological or social boundaries, and if so, how and why are those boundaries important in human life? What values are expressed by condemning, or allowing, incest? Does incest always manifest some vicious state of character? Would a rule allowing some incest prove harmful in the long run? Could anyone rightly reject a principle forbidding all incest?

What matters here is not how we answer the question about justification. It does not matter either the kind of argument that we give, or whether our answer is "yes" or "no." What matters is that *any* acceptable answer must point to some good, or some value, or some normative standard, in order to be an answer of the right kind. In that respect, an acceptable answer cannot simply point to the fact that we are inclined to condemn. To say simply that we are so inclined is inadequate, because the normative significance of that inclination is precisely what we are asking about. Saying that we are inclined tells us what we already know and does not help answer the question of whether judgments of condemnation are justified or not.

Now, either we can identify some value, or good, or normative standard that will allow us to answer the question about justification in the affirmative or we cannot. If we cannot, then we have not found good grounds for condemning incest, and we should probably stop doing so. And that is true regardless of anything science might tell us about the evolutionary origins of our inclination to condemn. For having asked for

a justification for condemning, we have been unable to find one. By itself that is enough to undermine our confidence that all incest is condemnable. After all, we have not been able to provide an account of why it is bad or to locate any bad-making features that show why all incest merits condemnation.

On the other hand, if we can provide an adequate answer to our question, then we have grounds for condemning all incest, and we should probably do so. And that is true regardless of anything science might tell us about the origins of our inclination to condemn. For the mere inclination to condemn is not the reason for our being justified in condemning. And our grounds for condemning—the goods/values/norms in light of which incest is shown to be bad—are not shown to be inadequate by any evolutionary story about our inclinations. If we have been able to offer an account of our good reasons for condemning incest, then the evolutionary story about our inclinations is neither here nor there.

Thus, however we answer the normative question about incest, the evolutionary account of our inclinations has no real part to play in providing the answer. And thus the science itself does not directly challenge (or support) anyone's values in the way that Greene suggests. All the normative work is done elsewhere. Of course, we might be able to come up with an answer to the normative question about incest (or another normative question) that is an answer of the right kind (in terms of some goods/values/norms) but that is not entirely convincing or decisive. Or there might be competing considerations, in terms of different goods/values/norms. And different answers might disagree over how serious a moral offense incest is (or is not). The point is not that justifications are an all-or-nothing affair. The point is that all the normative work must be done by whatever goods/values/norms are cited in answer to the normative question.

I conclude, then, that Greene's example of the direct route fails to show how cognitive science per se has any interesting moral implications. That said, the evolutionary account of our inclinations about incest might be what prompts us to ask the normative question in the first place. In that way, science would have relevance for normative ethics. But that is not Greene's direct route, nor is it an instance of cognitive science directly challenging anyone's values. Nor does it seem that science is necessary for raising the question. Faced with a case like Joe and Jane, but completely unaware of the evolutionary account, a person might reasonably wonder about whether we are right to condemn all incest—after all, don't we need good reasons to stand in the way of their love?

Greene notes that someone might object that "all the work" in his argument is done by the clause "and we have no additional reason to condemn their behavior" in his normative assumption. This is not exactly what I have argued, but it is close. Greene rejects this objection,

saying that this clause “simply closes off alternative escape routes, forcing a confrontation between the emotional response that is driving the judgment and the empirically based debunking argument that challenges it” (712 n. 72). This is mistaken. When we ask a normative question like the one Greene poses, the relevant challenge is to find some account that would justify us in answering that question one way or another. As we have seen, a story about the evolutionary origins of our inclinations is not necessary to raise the question, nor is it either necessary or sufficient to answer it.

III. IS THERE AN INDIRECT ROUTE?

As noted earlier, Greene’s indirect route is an updated argument for his much-discussed claim that cognitive (neuro)science provides support for consequentialism over deontology. The first premise of Greene’s new argument is the following claim:

The Central Tension Principle: Characteristically deontological judgments are preferentially supported by automatic emotional responses, while characteristically consequentialist judgments are preferentially supported by conscious reasoning and allied processes of cognitive control. (699)

The empirical support for this claim comes largely from “trolleyology,” in which persons are asked to judge whether a certain action is permissible. In the *switch* case, the action under consideration involves flipping a switch that will turn a runaway trolley away from a track with five people, thereby saving those five, and onto a track with one person, who will be killed by the diverted trolley. In the *footbridge* case, the action involves pushing a man from a footbridge onto the track below, in order to stop a runaway trolley and thereby prevent it from killing five people on the track. When asked if the action is permissible, most people say “yes” to *switch* and “no” to *footbridge*. This is the *switch-footbridge* effect.

What factors lead to the *switch-footbridge* effect? That is, what features are people responding to when they ponder trolley cases? According to Greene, there appear to be two key factors: “The first is whether the victim is harmed as a means or as a side effect, a factor that has long been cited by ethicists as relevant here and elsewhere. The second has to do with the ‘personalness’ of the harm. . . . The action in *footbridge* involves the application of *personal force*. That is, the agent directly impacts the victim with the force of his/her muscles” (709).

Greene regards “yes” in *switch* as a “characteristically consequentialist” judgment. Such judgments are “naturally justified in consequentialist terms (i.e., by impartial cost-benefit reasoning)” and “more diffi-

cult to justify in deontological terms because they conflict with our sense of people's rights, duties, and so on" (699). In contrast, saying "no" to *footbridge* is a "characteristically deontological" judgment. Such judgments are "naturally justified in deontological terms (in terms of rights, duties, etc.)" and "more difficult to justify in consequentialist terms, such as judgments against killing one person to save five others" (699). (Following Greene, I will drop "characteristically" and use an asterisk to denote these judgments.)

Based on empirical evidence, including fMRI imaging of the brains of people answering trolley questions, Greene concludes that deontological* judgments arise from automatic processes, whereas consequentialist* judgments arise from manual mode processes. When a person feels that "no" is the right thing to say in *footbridge*, that feeling comes from automatic processes. In contrast, the sense that "yes" is correct in *switch* is based, psychologically, in the kind of reasoning distinctive of manual mode. Thus Greene says, "the psychological *essence* of deontology lies with the automatic settings and the psychological *essence* of consequentialism lies with manual mode" (700).⁴

The Central Tension Principle (CTP) is an empirical claim. The second premise of Greene's argument is a normative claim:

The No Cognitive Miracles Principle: When we are dealing with unfamiliar* moral problems, we ought to rely less on automatic settings (automatic emotional responses) and more on manual mode (conscious, controlled reasoning), lest we bank on cognitive miracles.

Greene defines an unfamiliar* problem as one "with which we have inadequate evolutionary, cultural, or personal experience" (714). The argument for the No Cognitive Miracles Principle (NCMP) is based in the claim that automatic settings function well only when they have been shaped by trial-and-error experience. This could be personal experience, or the experience of our biological and cultural ancestors, passed on by genetic and cultural transmission. Since unfamiliar* problems are those for which we lack adequate experience, there is no reason to think that our automatic settings will function well when tackling such problems. Thus, "*it would be a cognitive miracle if we had reliably good moral instincts about unfamiliar* moral problems*" (715)—and we should not rely on such miracles.

4. My criticism of Greene does not focus on CTP. For challenges to Greene's empirical methodology and his interpretation of the data, see Guy Kahane, "On the Wrong Track: Process and Content in Moral Psychology," *Mind and Language* 27 (2012): 519–45.

To apply NCMP, we need to know which situations are unfamiliar*. So Greene adds a third premise. We can make an educated guess that a moral problem is unfamiliar* when it meets either of two conditions: (1) it is a problem that arises from recent cultural developments, or (2) it involves practical disagreement that does not appear to be based on disagreement over nonmoral facts.

By one or both of these criteria, Greene claims, *switch* and *footbridge* are unfamiliar*. That means that NCMP applies to them and cases like them, and that means we should rely less on automatic settings and more on manual mode, to avoid banking on cognitive miracles. From the first premise, CTP, we know that a deontological* judgment in a case like *footbridge* is preferentially supported by automatic settings—that is, we are relying on our automatic settings in coming to that judgment. On the other hand, when we make a consequentialist* judgment in such a case, we rely on manual mode. Thus we arrive at an important conclusion for normative ethics: in making a judgment in a case like *footbridge*, we should not rely on those intuitions that support our deontological* judgments. We should distrust the intuitions that push us toward deontology, because those intuitions, and the subsequent deontological judgments based upon them, arise from automatic processes, and NCMP tells us that we should not rely on manual mode in this situation. Thus cognitive science gives us “normatively significant information—information that can nudge us, if not propel us, toward new and interesting normative conclusions” (715). The indirect route “favors consequentialist approaches to moral problem solving, ones aimed solely at promoting good consequences, rather than deontological approaches aimed at figuring out who has which rights and duties, where these are regarded as constraints on the promotion of good consequences” (717).

Greene recognizes that some readers won't be satisfied. He imagines someone objecting that whatever is true of laypersons, surely nonconsequentialist philosophers engage in “manual mode”! They undertake plenty of conscious moral reasoning, and they base their deontological notions upon that reasoning. So it must be wrong to say that deontological* judgments are *only* supported by auto-mode intuitions and emotional responses.

In reply, Greene argues that nonconsequentialist philosophers are not primarily engaged in reasoning to figure out what's right or wrong. “Instead, their reasoning serves primarily to justify and organize their preexisting intuitive conclusions about what's right or wrong. In other words, what looks like moral rationalism is actually moral *rationalization*” (718). Crucially, in this rationalization, all the distinctively deontological content comes from intuitions generated by automatic settings.

To support this claim, Greene appeals to empirical work in which both philosophers and non-philosophers were presented with both *switchy*

and *footbridgesque* cases. Both groups tended to show a different pattern of judgment depending on which case was presented first. This reveals, Greene says, that both groups were engaged in a manual mode effort to be consistent. However, only the philosophers were also more likely to endorse the Doctrine of Double Effect (DDE), a deontological principle, if they were shown a *switchy* case first—a case that elicits an auto-mode response in support of a deontological* principle. This shows that the philosophers, but not the non-philosophers, “felt compelled to adjust their theoretical commitments to make them consistent with the judgments they’d made” (720). However, in doing so, the philosophers were not reasoning to new conclusions that changed the direction that their automatic settings were pushing them toward. On the contrary, the consistency that comes with DDE serves merely to codify results that are shaped by an auto-mode push toward “no” in *footbridgesque* cases. Thus Greene concludes: “Unless you’re prepared to say ‘yes’ to the *footbridge* case, your automatic settings are still running the show, and any manual adjustments that you’re willing to make are at their behest” (723). So even fancy deontological philosophizing ultimately depends on automatic settings, and hence violates NCMP.

How successful is Greene’s indirect route? Let us start with Greene’s claim that deontological philosophy is mere rationalization. As we’ve seen, Greene acknowledges that judgments in cases like *switch* and *footbridge* are sensitive to the distinction between harming as means and harming as side effect.⁵ This is a distinction that many philosophers consider morally relevant, and Greene supposes that such philosophers might invoke DDE as support for taking this distinction seriously. Such philosophers, Greene imagines, might also fault those who failed to invoke DDE to explain the difference between *switch* and *footbridge* cases. Greene then asks:

But why *should* they have invoked the DDE? Wherein lies its justificatory power? It’s been on the books for a long time, which gives it a dusty air of authority. But how did it get on the books? It seems that the DDE was codified because it was observed that certain intuitive patterns in moral judgment could be *summarized* by a set of principles now known as the DDE. We may infer this (inductively, not deductively) from the finding that lay moralists the world over make judgments consistent with the DDE while having no clue that they are doing so. This suggests, in other words, that the Doctrine doesn’t justify the judgments. Instead, *the judgments justify the Doctrine*. This evidence suggests that the justification for the DDE ultimately comes from nothing beyond the automatic settings that produce the pattern of judgment that it summarizes. (721)

5. Of course, they are not *only* sensitive to that distinction, according to Greene.

At this point, we should note an important difference between Greene's indirect route and his argument against deontology in previous work. In both his earlier and newer work, Greene maintains that deontology is an elaborate form of rationalization. But what Greene means by "rationalization" has shifted in an important way, even though he does not call attention to this shift. In an earlier essay, Greene explained the deontologist's rationalization with the following example: Alice goes on many dates and cites various factors to explain her likes and dislikes in romantic partners, factors such as intelligence, humor, and kindness. But investigation reveals that Alice always approves of partners over six-foot-four, and never approves of partners under that height. In this case, Greene concluded, we have good evidence that Alice "basically has a height fetish, and all of her talk about wit and charm and kindness is mere rationalization."⁶

Greene used the Alice example to show that we can identify a mere rationalization, in contrast to reasoning, by doing two things: (1) identifying a factor that predicts the rationalizer's judgments, and (2) showing that this factor is not plausibly related to factors that the rationalizer believes to be the basis of his or her judgments. Greene then suggested that the deontologist was like Alice. The factor that actually predicted deontological judgments in trolley cases was not anything deontologists themselves pointed to but instead whether the harming was "up close and personal."⁷ This was a factor that Greene took to be clearly morally irrelevant, and he reasonably expected deontologists to agree that such a factor was morally irrelevant, just as Alice might well agree that considerations of height should not exert the influence on her romantic judgments that they in fact do.

An important objection to Greene's earlier work was that, for all the experiments had shown, judgments in trolley cases might actually be sensitive not (merely) to whatever features make something "up close and personal," but (also) to other factors that are not morally irrelevant by the deontologist's own lights.⁸ Greene's view has evolved to deal with this objection. While Greene continues to maintain that deontology is rationalization, he now admits that one factor influencing judgments in trolleyology is whether the victim is harmed as a means or as a side effect—a factor that deontologists point to, and that we obviously cannot assume to be morally irrelevant without begging the question against the deontologist. This shift is crucial, because it means that Greene can no longer

6. Greene, "Secret Joke," 67.

7. *Ibid.*, 48.

8. Berker, "Normative Insignificance," 324; Kamm, "Neuroscience and Moral Reasoning," 334–35. For criticism of Greene's comparison of the deontologist with Alice, see Dean, "Does Neuroscience Undermine Deontological Theory?," 47–53.

claim that the deontologist is a rationalizer just like Alice. The rationalization that Alice gives is a kind of confabulation about her particular romantic judgments. What is really shaping her judgments is her (unrecognized) sensitivity to considerations of height, but in explaining her judgments she (unwittingly) makes up a story that refers to different considerations. If the deontologist's particular judgments in trolley cases were sensitive merely to considerations that make something up close and personal, then the deontologist would be like Alice.⁹ This would not show that deontological judgments were wrong, but it would show that the deontologist was not doing what she thought she was doing in making these judgments. It would reveal her to be a stranger to herself in an important way.

However, in Greene's updated argument, the deontologist's rationalization is something quite different: not confabulation relating to the factors behind specific judgments, but "*intuition-chasing*, conforming general principles to specific judgments that (mostly) follow the ups and downs of intuition" (718–19). This is a very different kind of "rationalization" than the confabulation illustrated by the example of Alice. And this difference is crucial to Greene's claims about the significance of cognitive science for ethics. To see why this difference matters, return to Greene's question about DDE: "Wherein lies its justificatory power?" This, I believe, is a perfectly good question. It is important to be clear about how the question arises. It does so as follows: faced with a couple of moral scenarios, we ask what to do, and we have an intuition that we should say "yes" to *switch* and "no" to *footbridge*. That is how we are inclined to answer. We then wonder whether this intuition actually corresponds to morally relevant features of the situation—that is, is this intuition tracking any good reasons? Someone then suggests that there is often a morally relevant difference between directly intending harm to another, as a means to some further aim, and merely foreseeing that one's action will cause harm as an unintended side effect. Call this the *i/f* distinction. The notion that this distinction is often morally relevant for evaluating actions can be stated as a general principle. Call this basic idea the Doctrine of Double Effect (DDE). Here are two versions of DDE:

To put things in the most general way, we should say that [DDE] distinguishes between agency in which harm comes to some victims, at least in part, from the agent's deliberately involving them in something in order to further his purpose precisely by way of their being so involved (agency in which they figure as *intentional objects*)

9. The same would be true if the deontologist's judgments were sensitive merely to what Greene now calls "personal force." The shift from "up close and personal" to "personal force" does not matter here.

and harmful agency in which nothing is in that way intended for the victims or what is intended does not contribute to their harm. Let us call the first kind of agency in the production of harm *direct*, and the second kind *indirect*. According to this version of the doctrine, we need, *ceteris paribus*, a stronger case to justify harmful direct agency than to justify equally harmful indirect agency.¹⁰

The DDE is the thesis that there is normally a stronger reason against an act if that act has a bad state of affairs (like an innocent person's death) as one of its intended effects than if that bad state of affairs is merely one of the act's unintended effects. In short, according to the DDE, it is harder to justify an act that has a bad effect if that effect is intended than if it is not intended.¹¹

All things equal, if a principle systematizes our intuitions about particular cases, that would seem to count in its favor. But even if we have a version of DDE that captures our intuitions about cases, we can continue to ask, as Greene does, about its justification: Is DDE really a good guide for decision making? That is, are we right to regard the i/f distinction as morally significant? This is Greene's question about justificatory power.

Since this question is about justification, an answer needs to identify some grounds for taking seriously the moral distinction that DDE points us to. We need some consideration—some good/value/norm—that reveals why we are correct (or not) to reason as DDE directs. Is there something important in human relationships that depends on respecting the i/f distinction? Is there something about our agency that explains why DDE is correct? Is there some value that is preserved/promoted/expressed when our attitudes and actions accord with the i/f distinction?

At this point, however, Greene supposes that the *only* answer we can give to the justificatory question is pointing to the fact that DDE is a good summary of our intuitions about different cases. And Greene believes—rightly, in my view—that if this is all we can say, then we have an unsatisfying defense of DDE. But what Greene overlooks is that this has nothing to do with automatic or manual mode settings. Rather the important point is that we can ask about the justification for DDE even after we have granted that DDE captures our intuitions about particular cases. For the question remains whether we have good grounds to take these intuitions seriously, to regard them as reliable guides. Once that question has been raised, it is unhelpful to reply merely that there is a principle (DDE) which tells us to judge in accordance with those intuitions, and

10. Warren Quinn, "Actions, Intentions, and Consequences: The Doctrine of Double Effect," in *Morality and Action* (Cambridge: Cambridge University Press, 1993), 175–93, 184–85.

11. Ralph Wedgwood, "Defending Double Effect," *Ratio* 24 (2011): 384–401, 384.

the support for which is that it summarizes our intuitions! That is pretty much begging the question, or at least leaving the circle of justification much too small.

So the “defense” of DDE that Greene imagines is indeed unsatisfying, but that can be seen without recourse to dual-process theory. Moreover, it might be that no better defense of DDE can be found. All we can say is that we have inclinations to judge in certain ways, and we can formulate a principle that summarizes them, but we cannot point to any good/value/norm that explains why reasoning this way is a good way of reasoning. We cannot say why we are right to observe the i/f distinction, or why DDE merits our adherence. In that case, we seem to lack a good response to the person who insists that the i/f distinction is morally irrelevant. And perhaps then we should reject DDE. But that has nothing to do with cognitive science.

On the other hand, suppose that we can offer a more compelling defense of DDE. Such a defense would reveal the good/value/norm in light of which we are justified in reasoning as DDE directs. Recent defenders of DDE have attempted to provide such defenses. Consider, for example, Warren Quinn’s account of DDE. Quinn points out that even if we formulate a version of DDE that does justice to our intuitions about cases, we face a further question about its rationale: “What, apart from its agreeing with our particular intuitions, can be said in favor of the doctrine? Indeed, why should we accept the intuitions that support it?”¹² Quinn answers that in cases of harmful direct agency, the relationship between the agent and the person harmed is different from the relationship in cases of harmful indirect agency. Both types of harmful agency might involve disrespect for those harmed, but with direct harmful agency the victim is “under our power and control in a distinctive way.” The agent of direct harm “has something in mind for his victims—he proposes to involve them in some circumstance that will be useful to him precisely because it involves them. He sees them as material to be strategically shaped by his agency.”¹³ This aspect of controlling others, of forcing them to play a role in one’s own plans, gives additional negative moral force to direct harmful agency, over and above the fact of harming or failing to prevent harm. What is so bad about bringing others under our power in this way? It violates an ideal of a human community as interaction among equals, in which each person “is to be treated, so far as possible, as existing only for purposes that he can share.”¹⁴ Overall, then, DDE merits being taken seriously because it captures a worthwhile ideal of human relationships: it

12. Quinn, “Actions, Intentions, and Consequences,” 176.

13. *Ibid.*, 190.

14. *Ibid.*, 192.

“embodies our sense that certain forms of forced strategic subordination are especially inappropriate among free and equal agents.”¹⁵

A different (but compatible) justification of DDE has been offered by Ralph Wedgwood, who claims that “DDE flows from a completely general feature of reasons for action: namely, in the significance for reasons for action of the agent’s degree of agential involvement in the consequences of the act.”¹⁶ According to Wedgwood, if the consequences of an act include a bad state of affairs, then the strength of the reasons against the act is determined not only by the badness of the state of affairs, but also the degree of “agential involvement” in bringing about the consequences. There are different types of agential involvement, and a person’s agency is more involved with a consequence that is intended than with a consequence that is unintended. Thus there are stronger reasons against actions that have a bad state of affairs among their intended, rather than merely foreseen, consequences. Thus, “the DDE is not just a widely-accepted idea that is supported by a wide range of intuitions; it can also be explained as flowing from a pervasive and fundamental feature of the normative domain.”¹⁷

Oddly, Greene never discusses these sorts of justifications. What matters for my purposes is not that either of these justifications of DDE is actually successful. What matters is that these justifications are not merely restating the fact that DDE captures our intuitions in particular cases. Nor are they adjusting the principle to fit new cases. Rather, they appeal to goods/values/norms that show why DDE is a principle that merits our adherence. And the key point is that *if* any such justification of DDE is successful, then we have an answer to Greene’s question about the justification of DDE. And in that case, dual-process theory is again irrelevant. For if we have an account of why DDE embodies a good way of judging and acting, then that account is not undermined by the fact that our automatic settings push us toward judgments that are consistent with DDE.

Of course, it might be that no such justification is successful. But then we are back to the situation we considered earlier—we have no answer to Quinn’s question, “What, apart from its agreeing with our particular intuitions, can be said in favor of the doctrine?” If we cannot answer that question, then indeed perhaps we should abandon DDE, or at least view it with considerable suspicion. But that is simply because we have been unable to give an adequate defense of DDE, and has nothing to do with cognitive science.

I conclude, then, that Greene’s indirect route does not support his conclusion that deontology as a school of moral thought is mere ratio-

15. *Ibid.*, 193.

16. Wedgwood, “Defending Double Effect,” 393.

17. *Ibid.*

nalization. That claim only looks plausible if we neglect the actual justifications that philosophers give for principles like DDE. I conclude also that Greene's indirect route fails to show how cognitive (neuro)science has normative implications, since the indirect route turns out to face a dilemma similar to the direct route. With the indirect route, cognitive science is meant to be normatively significant because it reveals that we should distrust deontological intuitions when confronting moral problems, with the result that our normative theories will tilt toward consequentialism. The indirect route gets going in response to a normative question like, "In solving moral problems, should we adhere to DDE? Is such a principle justified?"¹⁸ We can suppose that, in asking this question, we are already aware that we are *inclined* to judge in a way that tracks the i/f distinction. To answer this question, we need an account of why the i/f distinction is morally significant, if it is. Such an account must identify some good/value/norm that explains why we are correct to reason in accordance with the i/f distinction, rather than ignore it. If we cannot provide an account, then we have not been able to give good grounds for accepting DDE, and deontological moral theory is in a worse position—*regardless* of the dual-process theory of moral judgment. On the other hand, if we can give an account, then we are justified in accepting DDE—*regardless* of the dual-process theory of moral judgment. Either way, then, dual-process cognitive science does not matter in settling the normative question. If dual-process theory is relevant, it is because it gives us additional reason to ask for the justification. But even here cognitive science is not necessary to raise the question, as Quinn makes clear. Nor is dual-process theory either necessary or sufficient to answer it.

That said, cognitive science may be significant insofar as it reveals that our judgments about cases are *partly* influenced by factors that, by the judge's own lights, are morally irrelevant, such as "personal force." This would not show that the deontologist is a confabulator like Alice, nor would it show that deontological moral philosophy is mere intuition-chasing. But it would show that at least some of the factors behind our judgments are ones that we were unaware of, and that is a significant point about our (lack of) self-awareness.

IV. REPLY TO AN OBJECTION

Consider the following objection to my argument in the last section: "You claim what matters for normative ethics is the justification we can

18. For reasons of simplicity, I have followed Greene in focusing on DDE. Of course, some deontologists reject DDE, e.g., Thomas Scanlon, *Moral Dimensions* (Cambridge, MA: Harvard University Press, 2008). However, the basic shape of my critique of Greene can be endorsed by philosophers who think that the correct deontological account will appeal to a different principle, or set of principles, from DDE.

(or cannot) give for a principle like DDE, and that any adequate justification must identify some good/value/norm that supports the principle. But appealing to such justifications is out of bounds in this context. For surely our reasoning about goods/values/norms will be shaped by underlying psychological processes, including automatic processes. Thus what will *appear* to us as good or valuable or reasonable, and will hence *appear* to justify a deontological principle, will be the result of automatic processes. But it is precisely those processes that Greene's debunking argument shows to be unreliable guides for answering normative questions! Thus we cannot appeal to the results of reflective reasoning to show that the debunking has failed, because to do that would be to rely upon the very processes that the debunking argument shows to be unreliable."

I agree that, once the reliability of some belief-forming faculty has been cast into doubt, it is illegitimate to appeal to some particular results of that faculty to bolster our confidence in its reliability. But that is not what is going on here. The objection misconstrues the dialectic between Greene and the defender of deontology. The objection assumes that our automatic settings thoroughly shape the content of justifications like the ones that Quinn and Wedgwood offer. But nothing in the empirical data shows this. Indeed, Greene does not even claim this, since he never considers such justifications. And these justifications are not mere summaries of our intuitions about particular cases, but involve significant, new concepts (e.g., a community of interacting persons who act on reasons that others can share). Thus we have no reason to think that the results of reflective reasoning will be limited to what comes from automatic settings. Indeed, the deontologist might claim that her reasoning appeals to "philosophical intuition" (as opposed to "perceptual" or "dogmatic" intuition), which Greene does not view as suspect.¹⁹

19. See Greene, "Beyond Point-and-Shoot Morality," 724.