

## Criminalizing Endangerment - A Comment

Marcelo Ferrante

---

### Repository Citation

Marcelo Ferrante, *Criminalizing Endangerment - A Comment*, 65 La. L. Rev. (2005)  
Available at: <https://digitalcommons.law.lsu.edu/lalrev/vol65/iss3/3>

This Article is brought to you for free and open access by the Law Reviews and Journals at LSU Law Digital Commons. It has been accepted for inclusion in Louisiana Law Review by an authorized editor of LSU Law Digital Commons. For more information, please contact [kayla.reed@law.lsu.edu](mailto:kayla.reed@law.lsu.edu).

# Criminalizing Endangerment—A Comment

Marcelo Ferrante\*

We often criminalize types of actions, at least partially, because they cause or are likely to cause bad states of affairs, typically, harm to others. In his article *Criminalizing Endangerment*,<sup>1</sup> Professor Antony Duff proposes classifying offenses into two separate groups, *attacks* and *endangerments*, depending on whether or not offenders act with the intention to cause harm (or risk of harm).<sup>2</sup> Roughly stated, the rationale for the distinction between attacks and endangerments is that each kind of offense expresses a distinctive attitude toward the people or interests it harms (or threatens to harm). This distinction marks a significant moral difference in that each kind of offense instantiates a different kind of moral wrong. Professor Duff articulates this difference as follows: whereas attacks express “a practical hostility toward the interests or people” against whom they are addressed,<sup>3</sup> endangerments do not express such an attitude. At most, they manifest only a certain degree of indifference.<sup>4</sup>

The core of Professor Duff’s argument is best described as the conjunction of two separate theses. First, actions that equally cause (risk of) harm to others instantiate different kinds of moral wrongs, depending on whether or not such actions express an attitude of hostility. (I will call this the thesis of attitudes as wrong-makers.) Second, actions driven by the intention to cause (risk of) harm<sup>5</sup> always express hostility, whereas actions not so guided never express hostility. At most, they express merely indifference. (I will call this the perfect correspondence thesis.) The conjunction of these two theses yields that: 1) actions causing (risk of) harm that are performed with the intention to cause such harm (or risk thereof) and 2) actions that are not guided by such intention but are otherwise equal in effect to the former type of actions are instances of two

---

Copyright 2005, by R. A. Duff and Stuart Green.

\* Assistant Professor, Universidad Torcuato Di Tella School of Law (Buenos Aires, Argentina).

1. R. A. Duff, *Criminalizing Endangerment*, 65 La. L. Rev. 941 (2005).

2. See also R. A. Duff, *Criminal Attempts* 362–78 (1996) (proposing the same distinction).

3. Duff, *supra* note 1 at 943.

4. *Id.* at 944–45.

5. The term “intention,” as it figures in this thesis, does *not* encompass what is often called “oblique intention.” Duff, *supra* note 1 at 951. The term “harm,” in turn, refers to harmful states of affairs that are bad, all things considered. For instance, the surgeon who carries out an amputation that benefits the patient does not cause harm. *Id.* at 943.

different kinds of moral wrongs. The categories of attack and endangerment are meant to capture this distinction.<sup>6</sup>

If it is true that criminal laws should mirror the moral distinction between attacks and endangerments, then Professor Duff's argument, being correct, entails substantial implications for our current criminal law practices. For instance, current legal definitions of criminal offenses often disregard the contrast between causing harm with the intention to cause it and causing harm with mental states other than intention (e.g., knowledge or recklessness).<sup>7</sup> These criminal laws thus treat equally offenses that, according to Professor Duff's argument, constitute different kinds of moral wrongs. If, however, legal definitions of criminal offenses are to track moral differences, we should modify our current practices so that actions driven by the intention to harm and those not so driven are dealt with as offenses of different types.<sup>8</sup>

Similarly, current criminal laws often stretch the *mens rea* requirements for attempted liability to encompass not only actions in which agents intend to bring about harm, but also those in which agents only believe their conduct will cause harm.<sup>9</sup> Professor Duff's argument suggests that we should restrict the law of attempts to cases of intention and leave the rest, if they are to be criminalized at all, to some other form of inchoate liability.<sup>10</sup> Again, the problem would be that laws of attempts that have an overly broad scope blur the moral difference between attacks and endangerments.

I believe there are good reasons for granting Professor Duff's first thesis (the thesis of attitudes as wrong-makers). What I shall dispute in this comment is Professor Duff's second thesis (the perfect

---

6. Duff's use of the terms "attack" and "endangerment" assumes that the perfect correspondence thesis is true. "Attack" refers to harm-causing (or risk-imposing) actions guided by the intention to cause harm *and* that express hostility; "endangerment," in turn, refers to harm-causing (or risk-imposing) actions that are not driven by the intention to cause harm *and* do not express hostility. Since I contend that the perfect correspondence thesis is *not* true while conceding a version of the thesis of attitudes as wrong-makers, my use of the terms "attack" and "endangerment" will slightly diverge from Duff's. As I will henceforth use them, "attack" will denote an action that causes (risk of) harm and expresses hostility, while "endangerment" will mean an action that causes (risk of) harm but does not express hostility. Accordingly, I will be able to meaningfully ask whether harm-causing (or risk-imposing) actions driven by the intention to harm are always attacks, or whether harm-causing (or risk-imposing) actions that are not guided by the intention to harm could ever be attacks.

7. See, e.g., Model Penal Code § 210.2(1)(a) (Proposed Official Draft 1962) ("[C]riminal homicide constitutes murder when . . . it is committed purposely or knowingly.").

8. Duff, *supra* note 1 at 956–57.

9. See, e.g., Model Penal Code § 5.01(1)(b) (Proposed Official Draft 1962).

10. Duff, *supra* note 2 at 29–32, 362–78.

correspondence thesis), for I believe the alleged necessary correspondence between, on the one hand, intending harm and expressing hostility and, on the other hand, not intending harm and not expressing hostility (only, at most, indifference) does not hold. At the least, I believe it is more problematic than Professor Duff suggests.

Section I of this comment states why I think we may grant the thesis of attitudes as wrong-makers. Sections II and III advance my case against the perfect correspondence thesis.

### I. ATTITUDES AS WRONG-MAKERS

Actions instantiate wrongness by virtue of descriptive properties that they bear. We may call these properties *wrong-makers*, for bearing them makes an action wrongful. "Harm to others" is a wrong-maker: all other things being equal, the fact that an action of mine will harm you makes my action wrongful. I ought not to perform this action because it will harm you. Furthermore, harm to others is an uncontroversial wrong-maker. Claiming otherwise would commit us to an implausible view of morality under which the harmful consequences of our actions are not relevant considerations for deciding how to act.

In order for Duff's proposed moral distinction between attacks and endangerments to hold, it must be the case that besides performing actions that cause harm (or create risk of harm) to others, the attitudes agents manifest through these actions also count as wrong-makers. For, let us recall, the attitudes agents express toward those they harm (or risk harming) distinguish attacks from endangerments as wrongs of different kinds. This is in contrast to their effects or causal character, in which attacks and endangerments are indistinguishable. Thus again, what is peculiar to attacks is that they express hostility whereas endangerments do not, though they may express a degree of indifference.

In order to accommodate this distinction, we need a view of morality which entails not only preventing bad consequences and bringing about good ones, but also being sensitive to other forms of regard. So-called "expressive" theories of morality offer such a view.<sup>11</sup> These theories posit generally that according others their due is not just a matter of avoiding harm or contributing to well-being, but a question of holding and expressing through action the right sort

---

11. See, e.g., Elizabeth Anderson, *Value in Ethics and in Economics* chap. 2 (esp. at 33–34) (1993); Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. Pa. L. Rev. 1503, 1506–14 (2000); Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. Rev. 1659, 1681–85 (1992).

of attitudes. For instance, we owe our children more than a given level of physical or economic well-being—we ought to love them; caring for their welfare represents a part of this broader duty.

The duty not to harm others may be similarly analyzed. Our claim that we are not to be harmed by others is thus part of a more general claim that we be regarded as the kind of valuable beings we actually are. Actions which are equally harmful in their causal effects, then, can wrong their victims in different ways if they express different forms of regard by those who perform them. If, say, I harm you for the heck of it, I may wrong you in a different way than if I harm you inadvertently, even though both actions may be indistinguishable in terms of the risks they impose and the harms they actually bring about. In harming you for the heck of it, I express my view that the fact that you will be harmed gives me no reason at all not to perform the action which harms you; I straightforwardly deny that you have any value at all that may count against my whims—I may harm you for no particular reason at all.

In these terms, then, we may thus claim that in regarding you as such a valueless thing, I am violating your moral claim to be viewed as intrinsically valuable. If I harm you inadvertently, however, it is at most ambiguous whether my attitude toward you is as offensive as described above. I may fail to accord you the value that is your due, but I do not necessarily regard you as an utterly valueless thing that I may harm at will and for no particular reason.

Professor Duff's proposed terminology of *hostility* and *indifference* adequately characterizes the attitudes manifested in these examples:<sup>12</sup> In intentionally harming you, I express *hostility* toward you, for I straightforwardly deny you any value; in inadvertently harming you I reveal a sort of *indifference* toward you, for I simply do not care how or whether my behavior will affect you. A view of morality that distinguishes between such attitudes when evaluating the moral status of actions may lead to different judgments of wrongfulness for one and the other action. In such a view, each action instantiates a different kind of moral wrong by virtue of the attitudes it manifests, hence the use of different terms to designate these different wrongs: "attack" for the hostility-expressing wrong and "endangerment" for the wrong that imposes harm or a risk of harm without expressing hostility.

This sketchy argument shows why Professor Duff may plausibly claim that the wrongness of actions stems in part from the attitudes agents express in performing them, such that hostility-expressing actions which cause harm or risk of harm are wrongful in a way that

---

12. Duff, *supra* note 1 at 943–45, 947–48, 950–51.

otherwise equal but non-hostility-expressing actions are not.<sup>13</sup> To this extent, we may grant Duff's thesis of attitudes as wrong-makers.

## II. THE PERFECT CORRESPONDENCE THESIS

Duff's argument is completed by what I have called the perfect correspondence thesis. According to this thesis, the two kinds of moral wrongs captured by the terms "attacks" and "endangerments" are perfectly correlated with features of the actions' intentional structure. In particular, Duff contends, harm-causing (or risk-imposing) actions guided by the intention to cause (risk of) harm *necessarily* express hostility toward the persons or interests they are addressed to, whereas harm-causing (or risk-imposing) actions that are not guided by the intention to cause (risk of) harm *never* express hostility; in cases of recklessness or negligence, they may express some degree of indifference.<sup>14</sup>

If true, the perfect correspondence thesis offers a simple way of making our criminal laws reflect precisely the moral differences that the notions of attack and endangerment purport to capture. These concepts are, in Duff's terms, "thick" normative concepts: "concepts that reflect not just the causation of harm, but the way, the context, and the spirit in which harm is done."<sup>15</sup> Devising our criminal law rules to accurately reflect all the richness of the relevant moral factors might be an extremely burdensome task, if attainable at all. The perfect correspondence thesis, being true, solves this problem. It identifies a relatively "thin" factor, the intention to cause (risk of) harm, whose occurrence is invariably correlated with a richer factor, the expression of an attitude of hostility, and whose absence is invariably correlated with the absence of such a richer factor. We could thus have our criminal laws track the relevant, thick kinds of moral wrongs by avoiding clumping together, under the same legal types of offenses, actions guided by the intention to cause the relevant harm and those which are not thus guided—a fairly easy task.

There are good reasons, however, for doubting the perfect correspondence thesis. Harm-causing (or risk-imposing) actions guided by the intention to cause harm and otherwise equal actions that are not so guided may actually be understood as manifesting equivalent attitudes toward the interests or people they harm or threaten to harm. In particular, I will contrast actions performed with the intention to cause a harm (*H*) as a means to achieving a non-

---

13. *Id.* at 945–46.

14. *Id.* at 943–45.

15. *Id.* at 947.

harmful end (*E*) (which, according to the perfect correspondence thesis, are always attacks)<sup>16</sup> and otherwise equal actions, also performed with the intention to produce *E*, in which the agent foresees that *H* will obtain as a side effect (which, according to Duff, are always endangerments, no matter how certain the agent may be that her action will bring about *H*).<sup>17</sup> If, as I will be arguing, these actions do not necessarily express relevantly different attitudes toward the interests or people they harm, then it must be the case that they do not necessarily instantiate different kinds of moral wrongs by virtue of the attitudes they express. Actions that thus differ in their intentional structure but nevertheless express the same relevant attitudes toward harm must both be either attacks or endangerments, but this cannot be the case under the perfect correspondence thesis. My strategy will *not* be to deny an analytical distinction between intending harm as a means to a further end and foreseeing harm as a side effect. Rather, the point I will make is that what allows this distinction is unfit to mark a moral difference as significant as the one that the categories of "attack" and "endangerment" may be plausibly meant to capture.

Consider the following version of the examples Duff offers in support of his claim that, unlike intending harm as a means, foreseeing harm as a side effect does *not* express hostility.<sup>18</sup>

**Intention:** *A* wants to clear her view, which is blocked by her neighbor's tree. The tree stands near the border between the two properties but she cannot reach it to, say, cut it down. So, she lights a fire on her own property that will spread to part of her neighbor's property and, in its course, burn down the tree. *A* foresees that the fire will also spread in the opposite direction and burn down her own tree, which does not block her view.

**Foresight:** *B* wants to clear his view, which is blocked by a tree on his property. Due to where the tree stands, *B* cannot reach it to, say, cut it down. So, he lights a fire that will spread to where the tree stands and burn it down. *B* foresees that the fire will also spread in the opposite direction, reach his neighbor's property, and burn down a tree of his neighbor's, which does not block *B*'s view.

As these hypotheticals' titles suggest, only *A* intends to damage her neighbor's property; *B* just foresees that as a side effect of his

---

16. *Id.* at 950.

17. *Id.* at 951.

18. *Id.* at 947. I have modified the facts of Duff's examples in order to equalize all features other than the intent and foresight properties.

destroying his own tree, some damage to his neighbor's property will result. To say that *A* intends to damage her neighbor's tree is to say that the damage to the neighbor's tree plays a distinctive role in the practical reasoning that leads *A* to act as she does:<sup>19</sup> if *A* suspects that her original plan might fail to destroy her neighbor's tree, she is prepared to adjust her conduct so that she succeeds. For example, if, when she is about to start the fire according to her original plan, *A* realizes that unexpected wind conditions could prevent the fire from spreading far enough to reach her neighbor's tree (even though it would still reach her own tree), she will commit to adjusting her original plan, say, by waiting for more favorable weather conditions. This commitment is part of what it *means* that *A* intends to destroy her neighbor's tree.<sup>20</sup> Since, however, potential damage to *B*'s neighbor's tree does not play a comparable role in his practical reasoning, that is, he is not thus committed to adjusting his plan if equivalent circumstances arise, we may properly say that *B* does not intend to destroy his neighbor's tree; he just expects to destroy it as a side effect of burning down his own tree.

Having made clear what distinguishes these two cases, as regards their descriptive properties, we may now turn to the issue of possible moral significance. The challenge for the proponent of the perfect correspondence thesis is to show that the difference between the agents' practical reasoning reveals *A* to be hostile to her neighbor and *B* to be only indifferent toward his neighbor. However, claiming, as would Duff, that *B*, the agent in *Foresight*, is *indifferent* toward the destruction of his neighbor's tree mis-describes *B*'s attitude toward that event. In fact, from *B*'s perspective, destroying his neighbor's tree is, in fact, necessary to *B*'s own end of having a better view: if *B* must destroy his own tree in order to have a better view, then, given *B*'s beliefs about where he sets the fire and the causal laws governing how the fire spreads, the destruction of his neighbor's tree is required.

Since *B* possesses a positive attitude toward his desired end (having a better view), he must also have some sort of positive attitude toward those events that, according to his causal beliefs, are necessary occurrences on the way to his end. To suggest this positive attitude another way,<sup>21</sup> it would be good news for *B* that his neighbor's tree burns down because it would be evidence that *B*'s plan has succeeded and he now has his desired end: a better view. In this light, claiming that an agent has a positive attitude toward the

---

19. See Michael E. Bratman, *Intention, Plans, and Practical Reason* 140–43 (1999).

20. *But see id.* at 140, 191 n.6 (possible counterexample).

21. See Jonathan Bennett, *The Act Itself* 221–22 (1995).

occurrence of an event seems to exclude that the agent is indifferent toward the occurrence of that event. Rather, one is indifferent toward the occurrence of an event if one has neither a positive nor a negative attitude toward its occurrence.<sup>22</sup> Hence, my claim that the attitude manifested by the agent in *Foresight* is not indifference.

However, claiming that *B* is not indifferent toward the harm his action brings about is perhaps not yet to claim that the attitude he manifests in his action equals that expressed by *A* in *Intention*. One might still argue that attitudes like *B*'s, manifested in foreseeing harm as a side effect, differ relevantly from those manifested in intending harm as a means even though they are positive toward the harm. One might posit that foreseeing harm as a side effect is not as offensive toward the neighbor's interests as intending harm, for the following counterfactual holds: if *B* could improve his view without affecting his neighbor's tree, he might prefer that course of action rather than burning down his own and his neighbor's trees. More precisely, *B*'s attitude, as revealed in the original hypothetical, is consistent with his preferring to avoid harm, so we have no reason to deny he has that preference. Of course, it must then also be true that the foregoing counterfactual does not hold for *A*, the agent in *Intention*. And this is not the case. As captured by the counterfactual, *B*'s attitude does not differ from that of *A*. A similar counterfactual holds for *A* as well.<sup>23</sup> Let me explain.

As described above, *A* intends to destroy her neighbor's tree just as a means to having a better view. To say that an agent intends *H* just as a means to *E* is to state something about the agent's reasons for intending *H*, namely that the agent's only reason for intending *H* is that *H* contributes to the occurrence of *E*.<sup>24</sup> Whatever other properties *H* might have, they play no role in the agent's motivation to cause *H*. Accordingly, in *Intention*, *A*'s only reason for burning down her neighbor's tree is that it is necessary, under the circumstances, to achieving her end of having a better view. So, if the world were such that destroying her neighbor's tree was not necessary for *A* to have a better view, because, say, she could remove the tree without damaging it, she might prefer to do this rather than destroy her neighbor's property. Her positive attitude toward the intended harm, as revealed in the actual *Intention* scenario, is thus consistent with her preference for a possible world in which she has a better view and her neighbor's tree is not destroyed.

---

22. See Shelly Kagan, *The Limits of Morality* 170–71 (Oxford Univ. Press 1991) (having a positive attitude toward the occurrence of harm precludes being indifferent toward it).

23. See, e.g., Bennett, *supra* note 21 at 214–18; Horacio Spector, *Autonomy and Rights: The Moral Foundations of Liberalism* 109–11 (1992).

24. See, e.g., Kagan, *supra* note 22 at 130.

If sound, this argument indicates that the actions described in *Intention* and *Foresight* may result from the agents' equivalent dispositions concerning their respective neighbors. Both *A* and *B* appear to regard their neighbors' property interests in the same fashion. The attitude they manifest, call it hostility or something else, is a positive attitude toward harm based on (and only on) the perceived fact that *H* is necessary to achieving *E*. If *A*'s action is wrongful by virtue of the attitude it expresses, then *B*'s action must be equally wrongful. If *A*'s action and *B*'s action are thus equally wrongful by virtue of the attitudes they express, it is then not true that, all other things being equal, intending harm and not intending harm necessarily manifest relevantly different attitudes toward harm. It follows, that is to say, that the perfect correspondence thesis is false.

There may be, however, a possible limitation to the argument I have developed, based as it is upon the contrast between *Intention* and *Foresight*.<sup>25</sup> In those cases, it is assumed that both agents believe the harmful outcome of their actions will necessarily occur as a consequence of what they do. Under such conditions, the distinction between intending harm as a means and foreseeing harm as a side effect does not correlate with different attitudes toward the affected interests. Yet, this attitudinal symmetry may disappear if the agents believe that the harmful outcome is only a *likely* consequence of their actions. In effect, if *B* believes that the fire he starts in order to burn down his tree is only *likely* to spread and burn down his neighbor's tree, *B* could consistently *hope* that this likely outcome would not occur. If so, we cannot infer from *B*'s actions that he does not so hope. In contrast, the very fact that *A* *intends* to burn down her neighbor's tree, though by a means that makes such event only likely, seems to exclude the possibility that she can also hope the tree will not be destroyed. Thus, when the harmful outcome is believed to be uncertain, the action of the intending agent and the action of the foreseeing agent allow different inferences as to the attitudes they hold toward the interests they affect. While we can exclude the possibility that the intending agent hopes harm will not result from her action, we cannot exclude such a possibility in the case of the foreseeing agent. Articulating this distinction as the contrast between hostility and an attitude that does not amount to hostility, we can plausibly find in these cases the two different kinds of moral wrongs the notions of attack and endangerment are meant to capture.<sup>26</sup> If all

---

25. See Bennett, *supra* note 21 at 224–25.

26. Duff states that the same inference holds in cases where the foreseeing agent is certain of his causing the harmful outcome, and so claims that even in such cases, foreseeing harm as a side effect involves less offensive attitudes than

this were right, a qualified version of the perfect correspondence thesis may hold true in which not only intending harm but also foreseeing harm as a certain side effect are invariably correlated with the expression of hostility, whereas other mental states, including foreseeing harm as an uncertain side effect, would never express such an attitude. In any event, it would still be the case that Duff's (unqualified) perfect correspondence thesis does not hold.

### III. AIMING AT EVIL

Duff seems to concede that in focusing, as I have done so far, on the agents' dispositions to act in harmful ways, one may indeed conclude that, all other things being equal, there are no relevant attitudinal differences between intending and foreseeing agents.<sup>27</sup> But, Duff's argument goes, if we reject the position that criminal liability is to be grounded on these types of evaluations and focus on the character of the actions actually performed, we could notice a significant moral difference between actions guided by the intention to harm (even as a means to a further end) and those that are not so guided. The former type of actions, Duff writes, are "structured by the intention to [harm];"<sup>28</sup> they are "oriented toward and guided by the wrong that that intention involves."<sup>29</sup> The latter type of actions, in contrast, are "not thus structured by or oriented toward wrong."<sup>30</sup> To complete the argument, that is, to explain why in letting one's action being guided by harm one acts wrongly in a distinctive way,<sup>31</sup> Duff seems to resort to a well-known argument by Thomas Nagel:

[T]o aim at evil, even as a means, is to have one's action guided by evil. One must be prepared to adjust it to insure the production of evil: a falling-off in the level of the desired evil becomes a reason for altering what one does so that the evil is restored and maintained. But the essence of evil is that

---

intending harm as a means. Duff, *supra* note 1 at 951. Duff writes that the agent who foresees that a harmful event will ensue with certainty "might wish or hope that the harm would not ensue." *Id.* What does it mean, however, that the agent might "wish or hope" that the harm will not ensue when she knows it necessarily will ensue? This statement either makes no sense or refers only to the possibility of the agent's preferring a possible scenario in which achieving her end does not necessitate the harmful outcome. Yet, if Duff's statement refers to this second possibility, then it refers to something that fails to distinguish intending as a means from foreseeing as a side effect. My discussion of *Intention* and *Foresight* shows that both actions are equally consistent with the agents having such a preference.

27. *Id.* at 947-48.

28. *Id.* at 945.

29. *Id.* at 948.

30. *Id.*

31. *Id.* at 948 n.19.

it should *repel* us. If something is evil, our actions should be guided, if they are guided by it at all, toward its elimination rather than toward its maintenance. That is what evil *means*. So when we aim at evil we are swimming head-on against the normative current. Our action is guided by the goal at every point in the direction diametrically opposite to that in which the value of that goal points.<sup>32</sup>

In brief, if there is some value (*V*) and some thing (*T*) that instantiates *V*, the clearest mode of flouting *V*,<sup>33</sup> as far as *T* is concerned, is to intend the destruction of *T*.<sup>34</sup> For, in Nagel's terms, when we intend *T*'s destruction, we aim at evil and thereby "swim head-on against the normative current."<sup>35</sup> In contrast, by doing evil without aiming at evil, we still, presumably, flout *V*, though we do so less clearly and thereby swim against the normative current, but not *head-on* against it. Thus, because intending evil serves distinctively as the clearest mode of flouting the relevant value, it may serve as the basis for the proposition that intending evil, even as a means, expresses a particular kind of moral wrongness different in kind from (and more serious than) the wrongness of foreseeing evil as a side effect of aiming at a non-evil goal.

Under this interpretation, Duff's argument hinges crucially on the thesis that intending harm or evil is the clearest mode of flouting the pertinent value. Were this thesis false, the subsequent argument that finds in intending evil a distinctive kind of moral wrong should fall as well. The question is, then, whether the first thesis is correct. I contend that it is too narrowly framed. It leaves aside an important dimension of acting (im)morally. Once this dimension is properly accounted for, the case for intending evil as the clearest case of wrongdoing loses its ground.

Recall the facts of *Intention*. In that case, *A* intends to destroy her neighbor's tree. The state of affairs of the tree's being destroyed is bad, an instance of evil, let us say. The reason the destruction of the tree is evil is that it harms *A*'s neighbor. Since *A* intends to destroy her neighbor's tree and the tree's being destroyed is evil, if *A* aims at producing this state of affairs, it then follows that she aims at evil. However, the aspect of the intended state of affairs by virtue of which it is evil (i.e., that it harms the neighbor), although being part of what

32. Thomas Nagel, *The View from Nowhere* 181–82 (1986). Duff refers to a previous, but basically identical, version of it. See *supra* note 31 (citing Thomas Nagel, *The Limits of Objectivity*, in 1 Tanner Lectures on Human Values 75, 131–35 (Sterling M. McMurrin ed., 1980)).

33. I draw the concepts of *flouting values* and of there being different *modes* of flouting them from Robert Nozick, *Philosophical Explanations* 382–84 (1981).

34. Or, more accurately, intending *T* to stop being such that it instantiates *V*.

35. Nagel, *supra* note 32.

*A* intends, does not figure among *A*'s reasons for acting as she does. *Ex hypothesi*, *A*'s reason for burning down her neighbor's tree is that it will clear her view. True, this action will necessarily harm her neighbor, but this is not the reason why *A* burns the tree down. She burns down the tree *despite* (rather than *because*) it will harm her neighbor. This distinction suggests an even clearer case of flouting the value of the neighbor's interest in his or her tree: the case in which the agent intentionally burns down the tree *because* this will harm the neighbor. In this case, not only does the agent aim at evil, as *A* does in *Intention*, but she also takes what makes the intended state of affairs evil (i.e., that it harms the neighbor) as her reason (or one of her reasons) for bringing it about. We may say, using Nozick's terms, that in the latter case the agent intentionally does evil *qua* evil: her action does not merely coincide with wrongness, it *tracks* wrongness.<sup>36</sup> In *Intention*, however, *A* also does evil intentionally but her action does not track wrongness, it just fits wrongness.

In Kantian ethics, the moral action is such that the reason why the agent does it and the reason why the action is right are the same.<sup>37</sup> In acting morally in this strong sense, not only does the agent do what is right, she also connects herself with the value that makes her action right. If I do what is right, even though for a reason other than that by virtue of which my action is right, say, if I act for self-interested reasons where rightness stems from the fact that my action furthers others' interests, my action is still a right action even though it falls short of the moral ideal. Even though I do the right thing, I fail to connect myself to its pertinent value.

The contrast in the previous paragraph between pursuing wrongness *qua* wrongness and just having one's action fit wrongness involves a symmetric argument in the domain of acting immorally. When I intentionally do what is wrong and take what makes it wrong as my reason for doing it, not only do I act wrongly, I also connect myself with wrongness in the same way as the Kantian moral agent connects herself with value. If, in contrast, I intentionally do what is wrong without tracking wrongness, like the agent in *Intention*, my action is doubtless wrong but it falls short of the anti-ideal of immoral conduct just as the self-interested action of my previous example falls short of the moral ideal.

As Nagel claims in the above-quoted passage,<sup>38</sup> it follows from the very meaning of evil that we should not aim at it. Though true, this is only part of the story. It also follows from the concept of evil

---

36. See *supra* note 33 at ch. 4.

37. See Christine M. Korsgaard, *Creating the Kingdom of Ends* chap. 2 (1996).

38. Nagel, *supra* note 32.

that we should not take it as a reason for acting; we should not let our actions track wrongness. If sound, my point shows that aiming at evil is *not* the clearest mode of flouting the relevant value, for there is an even clearer case, to wit, aiming at evil *qua* evil. To return to Nagel's metaphor: only when we aim at evil *qua* evil do we swim head-on against the normative current; in merely aiming at evil (as the agent in *Intention* does) we certainly swim against the normative current, though not head-on against it. Thus, my point amounts to a rejection of Duff's argument for a distinction of moral wrongs that depends on whether the agent intends to cause harm or not for he grounds the difference on the distinctiveness of intending to cause harm as the clearest mode of flouting the relevant value.

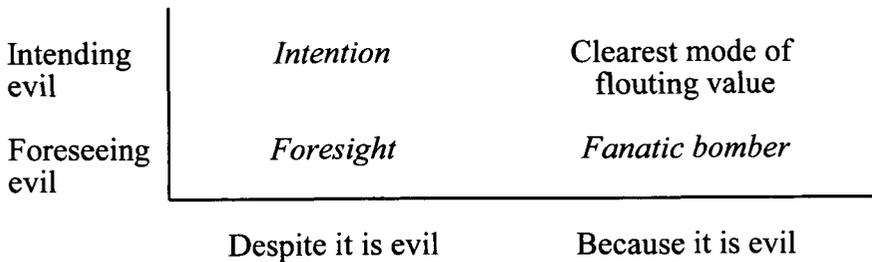
My point does not necessarily deny that, all other things being equal, intending harm is somehow more wrongful than foreseeing harm as a side effect. Indeed, my point is twofold: 1) For any case of wrongdoing in which the agent intends to cause a harm *H* without tracking wrongness, that is to say, without taking what makes that action wrong as her reason for doing it, there is a possible case in which the agent intends to cause *H* and tracks wrongness, that is to say, takes what makes the action wrong as a reason for doing it. 2) The latter case is the clearest case of flouting the value in relation to which the action is wrong. Accordingly, there is room for the argument that if the agent does not intend to cause *H* but only foresees it as a side effect, her action would be farther from the clearest mode of flouting the relevant value than the action guided by the intention to cause *H*. The action would still presumably flout that value, but in a less clear way.

Could this argument not ground Duff's proposal that we distinguish offenses by whether they are committed with the intention to cause harm or not? I think not. First, in order for Duff's proposal to be warranted along those lines it should be the case that, for any given evil state of affairs, intending it is always more wrongful than foreseeing it as a side effect. Second, even assuming that, all other things being equal, intending evil is more wrongful than foreseeing evil as a side effect, it will still not be the case that, for any given evil state of affairs, intending it is always more wrongful than foreseeing it as a side effect.

Take the following example. In a war driven by fanatic ethnic reasons, pilot (*P*) bombs some military facilities on the enemy's territory. *P* foresees some civilian casualties as a necessary side effect of the bombing. Even though *P* knows her action will kill civilians of the rival ethnicity, she does not intend to kill them; she aims to destroy the military facilities. *P* also believes, however, that all members of the rival ethnicity deserve to die and this belief has been crucial to her action: she would not be as willing to bomb the

enemy's military facilities were it to affect civilians of any other ethnicity. In other words, *P*'s action is wrongful partly by virtue of the fact that it will kill civilians of the rival ethnicity and this fact is part of *P*'s reasons for undertaking the bombing mission. *P*'s action of knowingly killing the civilians, then, is not guided by the intention to kill them, but it tracks the wrongness of killing them.

This example shows that foreseeing harm as a side effect and tracking the wrongness of the foreseen harm may appear together: one may knowingly do evil as a side effect *because* it is evil. In turn, the *Intention* scenario illustrates that an agent may intend to cause harm and still fail to track the wrongness of that harm; one may intend to do evil *despite* its being evil. On the other hand, the argument I am considering assumes that the clearest mode of flouting value involves *intending* evil *because* it is evil, therefore, all other things being equal, intending evil is closer to that clearest case of wrongdoing than simply foreseeing evil as a side effect. We should likewise grant that, all other things being equal, doing evil *because* it is evil is closer to the anti-ideal of immoral action than doing evil *despite* its being evil. As the cases of the fanatic bomber *P* and the *Intention* scenario show, intending evil (as opposed to foreseeing evil) and doing evil because it is evil (as opposed to despite its being evil) are independent properties: one may intend to do evil despite its being evil (as in *Intention*) and knowingly do evil because it is evil (as in the fanatic bomber case). The following graph displays these possibilities:



The vertical axis shows what we may call the *aiming dimension* of flouting the pertinent value. Along this dimension, intending evil ranks higher than foreseeing evil. The horizontal axis, in turn, displays what we may call the *tracking dimension*, along which doing evil because it is evil ranks higher than doing evil despite its being evil. The clearest mode of flouting the pertinent value stands on the upper-right region of the graph, implying that it involves the conjunction of the top positions along the two dimensions. On the lower-left region we find actions like that in *Foresight* in which the

agent knowingly does evil despite its being evil. Actions driven by the intention to do what is evil, like the action in *Intention*, are closer to the anti-ideal than those standing on the lower-left region of the graph. However, foreseeing evil as a side effect is no guarantee that foreseeing agents could never get closer to the anti-ideal. Cases like the fanatic bomber case show how this could occur.

In sum, once we allow for differences along the tracking dimension of wrongness, we cannot claim, as Duff's proposal requires, that, for any given evil state of affairs, intending it is always distinctively more wrongful or closer to the anti-ideal of wrongfulness than foreseeing it as a side effect.<sup>39</sup> Duff's proposal, then, loses its ground.

#### IV. CONCLUSION

Duff's argument would lead us to devise our criminal laws so that harm-causing (or risk-imposing) offenses are treated differently depending on whether they are committed with the intention to cause (risk of) harm or not. The reason Duff advances in support of this theory is that by intending harm, or in intending harm, offenders manifest relevantly different attitudes toward the victims they harm or impose risk upon. These differential attitudes, he argues, do not just show offenders to be differentially culpable. Victims are differentially *wronged* as a result. In section I, I sketched an argument that may account for this latter thesis. I pointed out a plausible view of morality under which actions may be wrongful—they may wrong others—not only by virtue of the evil consequences they impose on others, but also because of more subtle forms of regard that harmful actions may reveal, particularly, on account of the attitudinal dispositions they express.

In accordance with this view, I considered in section II whether the intention/non-intention divide tracks such attitudinal differences. I concluded that it does not because causing harm with the intention to cause it and knowingly causing harm may manifest equivalent attitudinal dispositions toward harm. Then, in section III, I considered the possibility that intending to harm when one *aims* at the production of evil could mark the pertinent moral difference. I argued that the view that could warrant such a difference fails to pay heed to an important dimension of wrongness. Once that dimension is properly taken into account, the distinctiveness of intending harm, as a particular and more serious kind of moral wrong vanishes.

---

39. Unless an argument is made that differences along the aiming dimension impact more heavily on the overall assessment of wrongness than those along the tracking dimension—an argument, by the way, that Duff does not make.

If my arguments are sound, we may have reasons for treating harm-causing (or risk-imposing) offenses as different kinds of wrongs depending on the attitudes offenders manifest through their actions. Yet, it will not follow that we should make intending to cause (risk of) harm the criterion by which the relevant classes of cases are distinguished.