KEN LEVY

THE SOLUTION TO THE PROBLEM OF OUTCOME LUCK: WHY HARM IS JUST AS PUNISHABLE AS THE WRONGFUL ACTION THAT CAUSES IT

I. INTRODUCTION

Many, and perhaps most, of us share the following two pre-philosophical intuitions:

(1) All else being equal, an intentional killer deserves harsher punishment than an agent who attempts to kill but fails.[1]

(2) All else being equal, an agent who unintentionally kills deserves harsher punishment than an agent who acts in precisely the same negligent or reckless manner and does not kill.[2]

---

[1] The Model Penal Code is consistent with (1). While it generally prescribes punishing "inchoate" offenses equally with their corresponding completed offenses, a position that is counter to the trend of most jurisdictions (see Paul H. Robinson and John M. Darley, *Justice, Liability, and Blame: Community Views and the Criminal law*, Boulder, Colorado and Oxford: Westview Press, 1995, p. 14), it does make an exception in § 5.05(1) for what it regards as the most serious offenses: "An attempt, solicitation or conspiracy to commit a [capital crime or a] felony of the first degree is a felony of the second degree." Incidentally, when I speak of intentionally killing, I shall be assuming that a non-exculpatory motive like revenge, jealousy, anger, sadism, or greed underlies the intention.

[2] There is yet a third intuition with which (1) and (2) should not be confused: an intentional killer deserves harsher punishment than an agent who unintentionally (negligently or recklessly) kills. I will not discuss this intuition in my paper because it leads to very different questions than do (1) and (2). They are: (a) Even if we concede that the intentional killer has worse character than the unintentional killer, why should this fact make the former deserve greater punishment? (b) Why should we think that the intentional killer has worse character than the unintentional killer in the first place? (c) Why do we think that even a *remorseful* intentional killer should still be punished more than a *remorseless* unintentional killer?

(1) and (2) significantly overlap. Both propositions state that when two agents with identical intentions perform identical actions in identical (relevant) circumstances and still produce different results, the agent who produces the greater harm deserves harsher punishment.

The problem with (1) and (2) is that they both run into a very powerful argument – what I shall refer to as the "Equal Punishment Argument". The Equal Punishment Argument suggests that agents deserve punishment only for what they have control over and that agents have control over only their bodily motions and intentions, not the external consequences of these bodily motions and intentions.[3] Therefore two agents who perform the same actions with the same intentions should be punished equally even if one of the agent's actions leads to harm and the other does not. Like (1) and (2) above, this argument is quite commonly discussed – and accepted – in the moral luck literature.[4] I too shall discuss it. But I shall not accept it. On the contrary, I shall argue that it is fundamentally flawed, that even if the external harm caused by one's action is out of her control, it should still be factored into her punishment.

While the justification for this point is actually quite simple, it has been surprisingly overlooked in the vast literature that has

---

[3] For the sake of clarity, I shall generally represent this distinction between bodily motions on the one hand and the harm that they produce on the other as a distinction between *actions* and their external *consequences*. So my use of the word *action* is meant to exclude whatever consequences the bodily motions produce outside the body. I make this point because we do not always use *action* in this way. For example, we sometimes characterize the agent's action as *killing* even though the death of the victim, which is necessary to make the agent's action a killing, constitutes an external consequence of the agent's bodily motions.

[4] Philosophers who accept the Equal Punishment Argument may not share intuitions (1) and (2). But they do not share these intuitions most likely because they have been *corrupted* by philosophical argument. That is why I use the term *pre-philosophical* in the first sentence of this paper. In the end, I would be very surprised if more than a handful of *laypersons* did not share intuitions (1) and (2). Paul H. Robinson and John M. Darley (*Justice, Liability, and Blame: Community Views and the Criminal Law* (Boulder, Colorado and Oxford: Westview Press, 1995, pp. 23, 25–28, 34–35, 41, 51, 159, 182–184, 198) bear me out on this point.

developed around this issue. The justification for this point is nothing more than the moral and legal concept, *assumption of risk*. I assume the risk of a given consequence "C" when I voluntarily perform a given action "A" and I can be reasonably expected to have known that A would (significantly) increase the probability of C. The paradigmatic example of this concept is the gambler at the roulette wheel, who assumes the risk of losing her money simply by placing it on a number before the wheel is spun. Precisely because she assumes the risk of losing her money in this way, she deserves whatever monetary fate the roulette wheel delivers. And this is the case even though the operations of the roulette wheel are out of her control. For by the conventions of gambling, when the gambler puts her money down on the roulette wheel, she is making a kind of "deal" with the casino. She is in effect saying that she is agreeing to let these events outside her control – i.e., "metaphysical luck"[5] – determine whether the casino will be entitled to take her money away or obligated to add more to it. And she makes this deal in the hopes that metaphysical luck will "go her way" and thereby lead to the latter rather than the former result.

Likewise, even though a given shooter (say) may not have control over what happens to the bullet once it leaves her gun, if she can be reasonably expected to know that her act of shooting will significantly increase the risk of certain harms (injury or death) and if she can be reasonably expected to know that these reasonably foreseeable harms are morally undesirable, then she is making a sort of "deal" with "the gods". Just as the gambler agrees with the casino dealer to let the *profit* status of her bet at the roulette wheel be determined by metaphysical luck, the risk-creating shooter is agreeing with another kind of dealer – the dealer of *morality* – to let the *moral* status of her shooting be determined in part by which of these reasonably foreseeable consequences ensues from her act of shooting. As a result, even though what happens to the shooter's target after the bullet leaves her gun is entirely outside the shooter's control, it still plays an essential role in determining the moral status of her act of

---

[5] I choose to use "metaphysical luck" rather than "moral luck" because I believe that the former has wider application.

shooting. If the shooter actually kills her victim, this consequence will make her act of shooting morally worse than if the shooter "only" injures or even misses her victim. And because the level of punishment an offender deserves generally corresponds to the moral status of her action, it follows that the level of the shooter's punishment should correspond to the moral status that the consequence of her act of shooting "bestows" upon her act of shooting. Contrary to the Equal Punishment Argument, then, it *is* indeed perfectly right and fair to factor the reasonably foreseeable harm produced by one's action, however out of her control it may be, into her punishment.

II. THE EQUAL PUNISHMENT ARGUMENT

For the sake of convenience, I shall refer to the intentional killer in (1) as "Intentional", to the unintentional killer in (2) as "Unintentional", to the agent in (1) who attempts but fails to kill as "Failed", and to the agent in (2) who acts just like Unintentional but does not actually kill anybody as "Lucky". I shall also refer to Intentional's victim in (1) as "Victim$_{IN}$", to Unintentional's victim in (2) as "Victim$_{UN}$", and to Failed's target in (1) as "Survivor".

In this section, I shall spell out what I take to be the strongest possible argument against (1) and (2) – what I shall refer to as the "Equal Punishment Argument". Sverdlik (1988) refers to the conclusion of the Equal Punishment Argument – step (9) below – as the "equivalence theory",[6] to proponents of the equivalence theory as "equivalence theorists", and to opponents of the equivalence theory as "non-equivalence theorists". In the course of my discussion, I will use Sverdlik's terms as well.

The Equal Punishment Argument attempts to show that if the only difference between two agents lies in the harm that they caused, and if this difference is attributable only to a difference in

---

[6] Michael S. Moore ("The Independent Moral Significance of Wrong-doing" *Journal of Contemporary Legal Issues* (1994): 237–281) refers to the equivalence theory as the "standard educated view".

what is outside their control, then they deserve to be punished equally.[7] While the Equal Punishment Argument applies equally to (1) and (2), each proposition involves a different kind of hypothetical situation. (1) involves something like the following:

---

[7] Those who subscribe to the Equal Punishment Argument (or at least to the conclusion of the Equal Punishment Argument – i.e., to the equivalence theory) include: Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30; "Insufficient Concern: A Unified Conception of Criminal Culpability" *California Law Review* 88 (2000): 931–953), Andrew Ashworth ("Criminal Attempts and the Role of Resulting Harm Under the Code, and in the Common Law" *Rutgers Law Journal* 19 (1988): 725–772), Lawrence C. Becker (Criminal Attempt and the Theory of the Law of Crimes" *Philosophy and Public Affairs* 3 (1974): 262–294), Joel Feinberg ("Equal Punishment for Failed Attempts: Some Bad But Instructive Arguments Against It" *Arizona Law Review* 37 (1995): 117–133), James J. Gobert ("The Fortuity of Consequence" *Criminal Law Forum* 4 (1993): 1–46), H.L.A. Hart (*Punishment and Responsibility* (New York and Oxford: Oxford University Press, 1968)), Sanford Kadish ("The Criminal Law and the Luck of the Draw" *Journal of Criminal Law and Criminology* 84 (1994): 1501–1523), Kimberly D. Kessler ("The Role of Luck in the Criminal Law" *University of Pennsylvania Law Review* 142 (1994): 2183–2237), Stephen J. Morse ("The Moral Metaphysics of Causation and Results" *California Law Review* 88 (2000): 879–894), Richard Parker ("Blame, Punishment, and the Role of Result" *American Philosophical Quarterly* 21 (1984): 269–276), J.C. Smith ("The Element of Chance in Criminal Liability". *The Criminal Law Review* (1971): 63–75), Steven Sverdlik ("Crime and Moral Luck" *American Philosophical Quarterly* 25 (1988): 79–86), Michael J. Zimmerman ("Luck and Moral Responsibility" *Ethics* 97 (1987): 374–386; "Taking Luck Seriously" *Journal of Philosophy* 99 (2002): 553–576, pp. 560–562).

Those who reject the Equal Punishment Argument and embrace instead the non-equivalence theory include: Judith Andre ("Nagel, Williams and Moral Luck" *Analysis* 43 (1983): 202–207), Brynmor Browne ("A Solution to the Problem of Moral Luck" *Philosophical Quarterly* 42 (1992): 345–356), Margaret Urban Coyne/ Margaret Urban Walker ("Moral Luck?" *Journal of Value Inquiry* 19 (1985): 319–325, "Moral Luck and the Virtues of Impure Agency" *Metaphilosophy* (1991): 14–27), Michael Davis ("Why Attempts Deserve Less Punishment than Complete Crimes" *Law and Philosophy* 5 (1986): 1–32), Antony Duff ("Auctions, Lotteries, and the Punishment of Attempts" *Law and Philosophy* 9 (1990): 1–37), John Greco ("A Second Paradox Concerning Responsibility and Luck" *Metaphilosophy* 26 (1995): 81–96), Barbara Herman ("Feinberg on Luck and Failed Attempts" *Arizona Law Review* 37 (1995): 143–149), Leo

(S1) Both Intentional and Failed aim their loaded guns at $Victim_{IN}$ and Survivor respectively and pull their triggers with the intention of killing their respective targets. They intend and do the very same thing in very similar circumstances. The result is this: while $Victim_{IN}$ dies immediately from

---

Footnote 7 (continued)

Katz ("Why the Successful Assassin Is More Wicked than the Unsuccessful One" *California Law Review* 88 (2000): 791–812), David Lewis ("The Punishment that Leaves Something to Chance" *Philosophy and Public Affairs* 18 (1989): 53–67), Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281), Ken O'Day ("Some Thoughts on Joel Feinberg's Modest Proposal: Is It Really Such a Modest Proposal After All?" *Arizona Law Review* 37 (1995): 243–249), Brian Rosebury ("Moral Responsibility and 'Moral Luck'" *Philosophical Review* 104 (1995): 499–524). For a helpful and thorough synopsis of different arguments for the non-equivalence theory, see Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281).

Those who do not clearly fall into either camp include: Bjorn Burkhardt ("Is There a Rational Justification for Punishing an Accomplished Crime more Severely Than an Attempted Crime?" *Brigham Young University Law Review* (1986): 553–571) (the non-equivalence theory is true but rationally unjustifiable), Gerald Dworkin and David Blumenfeld ("Punishment for Intentions" *Mind* 75 (1966): 396–404, pp. 396–399, p. 404) (neither retributivism nor utilitarianism can justify different punishment, in which case "either the theories are unsatisfactory or our practices should be changed"), Henning Jensen ("Morality and Luck" *Philosophy* 59 (1984): 323–330) (the Equal Punishment Argument is true with respect to (1) but false with respect to (2)), Norwin Richards ("Luck and Desert" *Mind* 95 (1986): 198–209) (the Equal Punishment Argument would be correct if, contrary to fact, we had omniscient knowledge of people's characters and therefore their just deserts), Stephen J. Schulhofer ("Harm and Punishment: A Critique of Emphasis on the Results of Conduct in the Criminal Law" *University of Pennsylvania Law Review* 122 (1974): 1497–1607) (some justifications for factoring harm into punishment work but only in a limited number of situations under certain assumptions).

For a critical but, in my opinion, opaque response to Norwin Richards ("Luck and Desert" *Mind* 95 (1986): 198–209), see Jonathan Adler ("Luckless Desert is Different Desert" *Mind* 96 (1987): 247–249). The reader should be aware that I leave out any discussion of Bernard Williams ("Moral Luck" *Aristotelian Society* Supp. 50 (1976): 115–136) because, despite the title, the particular content of his paper lies rather far afield from the issues that I discuss in this paper.

Intentional's bullet, the bullet intended for Survivor is intercepted by a passing bird.[8]

And (2) involves something like this:

(S2) While both Unintentional and Lucky are driving down different streets that are very similarly situated, they look away from the road for several seconds at a plane flying overhead. During the time that Unintentional and Lucky look away from the road, Victim$_{UN}$ walks in front of Unintentional's car and nobody walks in front of Lucky's car. In the end, while Unintentional hits and kills Victim$_{UN}$, Lucky does not hit or kill anybody.

Given (S1) and (S2), the Equal Punishment Argument proceeds as follows:

(3) An agent deserves punishment only for that over which she has control, not for that over which she has no control. And the severity of punishment that she deserves correlates with the moral magnitude of that over which she has control.[9]

---

[8] Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30, pp. 8–12) cleverly puts Intentional and Failed (among others) into the *very same* situation rather than into two isolated situations and then argues that our intuitions about how Intentional and Failed should be treated run in the very opposite direction than (1). But even if Alexander's assessment of our intuition about this new situation is correct, this point hardly undermines (1). For our intuition about this new situation may be influenced by the now different fact that Intentional and Failed are acting more or less together, as a group rather than just as individuals.

[9] Margaret Urban Walker ("Moral Luck and the Virtues of Impure Agency" *Metaphilosophy* (1991): 14–27, p. 16) refers to (3) as the "control condition" – "the intuitive principle limiting moral assessment to just such factors as an agent controls". See also Thomas Nagel ("Moral Luck" *Aristotelian Society*, Supp. Vol. 50 (1976): 137–152. Reprinted in Thomas Nagel (*Mortal Questions* (Cambridge: Cambridge University Press, 1979)), 24–38, and Gary Watson ed. (*Free Will* (Oxford: Oxford University Press, 1982)), 174–186). Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30, 22 ff.) argues that the control condition is consistent with allowing certain kinds of moral luck, though not outcome luck, to affect an agent's punishability. Philosophers who reject (3) include: Judith Andre ("Nagel, Williams and Moral Luck" *Analysis* 43 (1983): 202–207), Brynmor Browne ("A Solution to the Problem of Moral Luck" *Philosophical Quarterly* 42 (1992): 345-356, pp. 350 ff.), Margaret Urban. Coyne/ Margaret Urban. Walker ("Moral Luck?" *Journal of Value Inquiry* 19 (1985): 319–325; "Moral Luck and the Virtues of Impure Agency" *Metaphilosophy* (1991): 14–27), Ken O'Day ("Some Thoughts on Joel Feinberg's Modest Proposal: Is It Really Such a Modest Proposal After All?" *Arizona Law Review* 37 (1995): 243–249). I too shall argue against (3) in Section VII.

(4) An agent should be given what she deserves – no more, no less.

(5) ∴ An agent should be punished only for what is in her control, not for that over which she has no control. And the severity of her punishment should correlate with the moral magnitude of that over which she has control.

(6) ∴ If the only difference between two agents regards something over which neither agent has control, if they are equal in all respects over which they do have control, then one agent should not be punished more than the other. Instead, they should be punished equally.

(7) In any given situation, an agent has control at most over (a) her state of mind and (b) her bodily motions. She does not have control over (c) what occurs outside her body.[10]

(8) ∴ Intentional and Failed have equal control over their intentions and their actions (pulling the trigger) and equal lack of control over what the bullets do once they leave their guns. Likewise, Unintentional and Lucky have equal control over their actions (looking at the plane while driving) and equal lack of control over whether or not a pedestrian crosses the path of their cars.[11]

(9) ∴ Even though Intentional's bullet kills $Victim_{IN}$ and Failed's bullet does not kill Survivor, and even though Unintentional kills $Victim_{UN}$ and Lucky kills nobody, both Intentional and Failed

---

[10] See Stephen J. Morse ("The Moral Metaphysics of Causation and Results" *California Law Review* 88 (2000): 879–894, p. 883).

[11] Both Intentional and Failed on the one hand and Unintentional and Lucky on the other must have equal control over their actions. If one agent has more control over her action than the other, then one might try to avert the problems raised by the Equal Punishment Argument by arguing that their different punishments are justified not by the different results but rather by their different levels of control. Since the debate between the equivalence theorist and the non-equivalence theorist is solely over whether or not a difference in harm alone should lead to a difference in punishment, we must keep everything that obtains up to and including the wrongful actions equal.

on the one hand and both Unintentional and Lucky on the other should be punished equally.[12-14]

Because (9) conflicts with (1) and (2), and because (1) and (2) represent two of our intuitions, (9) is counter-intuitive. So our task at this point is either to "bite the bullet" and drop (1) and (2) or to reject the Equal Punishment Argument. In sections VI and VII, I shall attempt to justify the latter option.

---

[12] Michael J. Zimmerman ("Luck and Moral Responsibility" *Ethics* 97 (1987): 374–386) argues that in addition to this conclusion, Unintentional and Intentional *should* indeed be punished at least to *some* extent. It follows, then, that Lucky and Failed should be punished to some extent as well. Zimmerman makes this point in order to avoid the possibility of arguing the other way around – i.e., of starting from the premise that Lucky and Failed *should not* be punished at all and thereby arriving at the conclusion that neither should Unintentional and Intentional.

[13] The Equal Punishment might equally put in terms of *risk*. The agents within each pair (i.e., Intentional and Failed on the one hand and Unintentional and Lucky on the other) equally posed an equal risk to others and equally can be reasonably expected to have avoided posing this risk. So far, then, their punishments should be equal. The only respect in which each agent is unequal – namely, the fact that the risk was realized in Intentional's and Unintentional's case and not in Failed's and Lucky's case – is precisely the respect in which their punishments may not differ. For Intentional and Unintentional had no more control over the fact that the risk was realized in their case than Failed and Lucky had over the fact that the risk was not realized in their case.

[14] Notice, the equivalence theorist and the non-equivalence theorist differ only with regard to the relative punishments that Intentional and Failed should receive. They do *not* necessarily differ with regard to the punishment that each deserves *independently* of the other. It is perfectly possible for a given equivalence theorist to think either (a) that Intentional should be punished to the same extent that a given non-equivalence theorist thinks that Intentional should be punished or (b) that Failed should be punished to the same extent that a given non-equivalence theorist thinks that Failed should be punished. But because they still do differ on the relative punishments Intentional and Failed deserve, the equivalence theorist cannot think *both* (a) *and* (b) (at least not with respect to the *same* non-equivalence theorist).

III. SOME CONSIDERATIONS ABOUT THE EQUAL PUNISHMENT ARGUMENT

A number of things need to be said about the Equal Punishment Argument. First, what is in question here between (1) and (2) on the one hand and the Equal Punishment Argument on the other is which approach we should take toward punishment – "action-extrinsic" or "action-intrinsic". On the one hand, the action-extrinsic approach suggests that the level of punishment for an offender should be determined at least in part by the level of harm that she causes – i.e., by her "wrongdoing".[15] This approach, which is embraced by the non-equivalence theorist, is supported by propositions (1) and (2). On the other hand, the action-intrinsic approach suggests that the level of punishment for an offender should be determined not by the consequences of her actions but only by her actions themselves independently of their consequences – i.e., by her "blameworthiness". The Equal Punishment Argument obviously follows the action-intrinsic approach. In the end, then, the reader may profit by remembering this basic distinction: while the motto of the equivalence theorist is "same action means same punishment" (regardless of action-consequences), the motto of the non-equivalence theorist is "different harm means different punishment" (even if the actions producing the different harms were identical).

Second, the Equal Punishment Argument says nothing about *fairness*. Yet it seems as though the Equal Punishment Argument could equally be formulated in terms of fairness rather than desert. If so, premise (3) would look more like this:

(3*) It is fair to punish an agent only for that over which she has control, not for that over which she has no control. And it is only fair that the severity of her punishment be determined by the moral magnitude of that over which she has control.

Would changing (3) in this way significantly change the overall argument? I do not think so. On the contrary, (3) and (3*) lie very

---

[15] I say *at least in part* rather than *only* because the action itself must also play a role in determining the appropriate level of punishment. For this reason, we should leave room for the possibility of punishing certain kinds of blameworthy actions that do not necessarily produce any harm at all – namely, actions falling into the categories of solicitation, conspiracy, accomplice liability, and attempt.

close together in conceptual space. There seems to be a close *explanatory* relation between them: (3*) seems to be the case because (3) is the case.[16] Of course, because the explanans and the explanandum cannot be identical, my notion that (3) explains (3*) assumes that the two premises are conceptually distinct. And one might argue that I have no reason to believe this, that I have no reason not to go even further and *identify* the two premises. But I do not believe that the two premises are identical with each other for the simple reason that it still seems possible for fairness and desert to come apart. It still seems possible, that is, for there to be (a) situations in which it would be unfair to give somebody what she deserves as well as (b) situations in which it would be fair to give somebody what she does not deserve.[17] Still, because these kinds of counterexamples are so rarely instantiated, it is safe to say that for the most part, the equivalence theorist may use both premises interchangeably. For this reason, I shall from time to time speak of the equivalence theorist's position in terms of fairness rather than in terms of desert. I do not believe that this terminological variation will beg any questions against, or unnecessarily weaken, the Equal Punishment Argument or the equivalence theory.

Third, one might wonder why I have framed the Equal Punishment Argument in terms of punishment rather than blame. One might argue that the same kind of point would fall out, that whether or not a given agent should be blamed and how much also depends on whether or not she had control over the consequences of her action and how much. I choose to put this argument in terms of punishment rather than blame, however, because I believe that it is our intuitions about *punishment* that really motivate the problem. If we frame the issue in terms of blame, then we may solve it simply by

---

[16] Of course, it is possible that the explanatory relation goes the *other* way, that (3) is the case because (3*) is the case. Either way, it is not important for the purposes of this discussion to figure out the correct direction of the explanatory relation.

[17] I shall discuss an example of (a) in Section VIII (Objections 4 and 5). Regarding (b), it seems clear enough that anyone guilty of a minor parking offense does not deserve ten days in jail. Yet one might very well argue that such punishment is still *fair* if the offender had ample notice of this possibility and everybody else guilty of the same offense is punished to the same degree.

concluding that we should blame one only to the extent that one's action itself violated a moral or social norm. In this case, Intentional and Failed on the one hand and Unintentional and Lucky on the other come out equal – i.e., equally blameworthy.[18] But this kind of solution does not work as well when the issue is framed in terms of punishment rather than blame. For even *after* the Equal Punishment Argument and even if we think that Intentional/Unintentional is *not* more blameworthy than Failed/Lucky, many of us *still* have the intuition that Intentional/Unintentional should be punished *more* than Failed/Lucky.[19] And this itself is an interesting result. What it suggests is that, according to our intuitions at least, "punishment desert" does not always directly correlate with blameworthiness, that there may be situations in which different punishment is consistent with equal blameworthiness.[20] Again, the other factor in addition to blameworthiness that may help to determine the appropriate level of punishment is the kind and extent of harm that the agent caused.

Fourth, in accordance with the second sentence in (3) above, I shall assume throughout this paper that the nature and degree of punishment that one receives should be proportional to her moral desert. Three points need to be made about this seemingly trivial assumption. First, it cannot be *entirely* trivial because not *all* philosophers agree with it. Some philosophers who subscribe to consequentialist justifications of punishment might argue that there are at least some situations in which the punishment *should not* "fit the crime"; situations in which the consequences of punishing an innocent person, absolving a guilty person, or giving a person either more or less punishment than she deserves would justify, if not require, such treatment. In this paper, I will simply not concern myself with these issues. Again, I will simply assume that the situations that I discuss are not these situations, that the agents in the

---

[18] This is the kind of solution that Judith Jarvis Thomson ("Morality and Bad Luck" *Metaphilosophy* 20 (1989): 203–221) proposes.

[19] I deliberately leave the expression "more blameworthy" ambiguous between deserving of more blame and more deserving of blame. Michael J. Zimmerman ("Luck and Moral Responsibility" *Ethics* 97 (1987): 374–386) offers this helpful distinction.

[20] Michael J. Zimmerman ("Taking Luck Seriously" *Journal of Philosophy* 99 (2002): 553–576, pp. 561–562) concedes the possibility of this point.

situations that I discuss should indeed receive the punishment corresponding to their level of desert. Second, this assumption hardly begs any questions against the equivalence theorist. For the equivalence theorist will most likely be perfectly happy to accept it as well. She will most likely *agree* with me that an agent's punishment should be proportional to her desert. What she and the non-equivalence theorist will instead disagree about is what a given agent deserves in the first place. Once again, while the equivalence theorist will argue that the agent's moral desert is, and therefore her punishment should be, determined only by her action, the non-equivalence theorist will argue that the agent's moral desert is, and therefore her punishment should be, determined not only by her action but also by the harmful consequences of her action. It is *this* disagreement that shall constitute the primary focus of my paper. Third, this assumption does not commit me to any particular theory about the *purpose(s)* of punishment. It is perfectly consistent with both forward-looking (consequentialist) as well as backward-looking (deontological or retributivist) theories. For one might think that punishment should be proportional to desert either for consequentialist reasons (e.g., such a principle maximizes deterrence and respect for the law) or for retributivist reasons (e.g., it is just a basic moral truth that one should get what one deserves). I shall have a bit more to say about the purpose(s) of punishment, however, in my upcoming reference to Duff and in Section VII.

Finally, it is important in a discussion like this to account for the intuitions that the Equal Punishment Argument challenges, to offer an explanation of why the Equal Punishment Argument strikes us as counter-intuitive. Otherwise, the Equal Punishment Argument gets a "free ride". It wins not on the merits but merely because the other side failed to show up for the game. Once again, the intuitions that the Equal Punishment Argument challenges are (1) and (2). And the best defense of these intuitions that I have been able to find is provided by Duff (1990, pp. 30–37). Duff argues that the amount of harm an agent causes *should* be factored into her punishment. For (a) the amount of harm that an agent causes *does* matter; (b) (a) *should* be communicated to the agent and to the wider community; and (c) the best, if not the only, way to accomplish (b) is by

factoring the amount of harm that the agent causes into her punishment.[21]

The equivalence theorist's natural objection to this point will be that while the amount of harm does of course matter, this does not mean that it should be reflected in the relative sentences Intentional and Failed (for example) receive. For the conclusion that we should factor harm into Intentional's and Failed's punishments is outweighed by two other considerations. First, it just is not fair to Intentional. (See the second consideration above.) Second, while the fact that Failed failed to kill Survivor is good news for Survivor, factoring this good news into Failed's punishment sends the message that this good news is to Failed's credit. But this message is false. It is not *because of* anything Failed did that Survivor was spared; it is entirely *in spite of* what Failed did that Survivor was spared. For both these reasons, the amount of harm that a given individual causes should *not* be factored into her punishment.

The equivalence theorist's second argument – the "message argument" – presupposes that punishment sends out messages only about blameworthiness, not about wrongdoing. But this assumption is incorrect. Punishment sends – or at least *should* send – out messages about *both* blameworthiness *and* wrongdoing. But the equivalence theorist's first argument – the "fairness argument" – is more difficult to dispose of. Indeed, what it shows is that Duff's proposal provides only half of the response necessary to undermine the Equal Punishment Argument. The half that is missing – the half that Duff does not address – is this fairness issue. One of the main tasks of my paper, then, is to complement Duff's argument above, to show that it is indeed fair to factor into one's punishment this harm that matters. I shall confront this task primarily in Sections VI and VIII.

---

[21] Notice, (b) uses what are arguably consequences (the various purposes for which (a) should be communicated to the wrongdoer and the wider community) to justify the factoring of harm into the wrongdoer's punishment. To this extent, Duff's position is consequentialist. Nigel Walker (*Why Punish?* (Oxford: Oxford University Press, 1991) pp. 79–82) makes a similar point about (b). See also Jonathan Jacobs ("Luck and Retribution" *Philosophy* 74 (1999): 535–555, p. 538, p. 552).

IV. METAPHYSICAL LUCK

I think that the proper response to the Equal Punishment Argument requires some critical clarifications regarding "metaphysical luck". I shall present these clarifications in this section.

By *metaphysical luck*, I mean whatever one lacks control over that is causally relevant to the outcome of her action.[22] Metaphysical luck in Intentional's case, for example, is whatever happens with the bullet and to Victim$_{IN}$ after Intentional pulls the trigger.

Whether or not metaphysical luck is "on Intentional's side" or "*not* on Intentional's side" depends on our standard of evaluation.[23] I can think of at least three such standards. First, the agent's own desires or intentions. By this standard, metaphysical luck is on only Intentional's side, not Failed's side. For while both wish to hit and kill their targets, only Intentional succeeds. Conversely, to another person (e.g., an inexperienced police officer) who does not wish to kill but accidentally does, metaphysical luck is *not* on her side. The second possible standard is how much punishment the agent will suffer if she is caught. If Intentional is caught, she is likely to be punished more harshly than she would have been punished had she failed to kill Victim$_{IN}$. In *this* respect, then, metaphysical luck is *not* on Intentional's side. Conversely, if Failed is caught for attempting to kill Survivor, we might say that metaphysical luck *is* on her side in the respect that she will receive less punishment than she would have had she succeeded in killing Survivor. The third possible standard is character assessment. If we consider Intentional to be a worse person than Failed because only Intentional, not Failed, is now a *murderer*, then metaphysical luck is more on Failed's side than on Intentional's side. For this difference in character assessment turns entirely on a difference in metaphysical luck. Conversely, if we consider them both

---

[22] See Jonathan Jacobs ("Luck and Retribution" *Philosophy* 74 (1999): 535–555, p. 539), Michael J. Zimmerman ("Taking Luck Seriously" *Journal of Philosophy* 99 (2002): 553–576 p. 559, n21).

[23] While the notion of luck's being "on one's side" may be somewhat anthropomorphic, I mean to suggest nothing more by this expression than in favor of one's preferences or interests.

to be equally reprehensible for equally attempting to kill, [24] then one might argue that metaphysical luck is more on Intentional's side than on Failed's. For Failed is being judged to be just as bad as a murderer even though she herself is not (let us assume).[25]

I was careful above to speak in terms of metaphysical *luck* being *on one's side* rather than in terms of "metaphysical *luckiness*" (or metaphysical luck *not* being on one's side rather than "metaphysical *un*luckiness"). For there is a distinction between the two kinds of expressions.[26] Again, when I say that metaphysical luck is on one's side, all I mean to suggest is that certain factors outside one's control helped to bring about the result either (a) that she wanted or that is more favorable to her in terms of either (b) the level of punishment she receives or (c) character assessment. But this point is distinct from saying that the agent is *metaphysically lucky*. For to say that a given agent is *metaphysically lucky* generally means not only that metaphysical luck is on her side but also that the factors outside her control which helped to bring about the result in accord with her preferences or interests were "out of the ordinary" – unusual, atypical, unexpected, surprising, improbable, or abnormal.[27] So whether or not metaphysical luck is on her side, Intentional in (S1) is *not* metaphysically lucky. For the factors that helped to bring about the result that she wanted once the bullet left her gun were perfectly *normal* – e.g., the gun did not misfire, the laws of nature continued to operate as they did before she pulled the trigger, nothing suddenly interfered with the path of the bullet, and the bullet hit in a vital spot. But, assuming that metaphysical luck was indeed on Intentional's side, Intentional *would* have been metaphysically lucky if she had

---

[24] See Gerald Dworkin and David Blumenfeld ("Punishment for Intentions" *Mind* 75 (1966): 396–404, p. 397), Gregory S. Kavka ("Some Paradoxes of Deterrence" *Journal of Philosophy* 75 (1978): 285–302, p. 289).

[25] See Peter Winch (*Ethics and Action* (London: Routledge and Kegan Paul, 1972), pp. 149–150).

[26] See Barbara Herman ("Feinberg on Luck and Failed Attempts" *Arizona Law Review* 37 (1995): 143–149, pp. 147–148).

[27] Naturally, then, to say that an agent is metaphysically *un*lucky is to say that metaphysical luck is *not* on her side *and* that the factors outside her control which helped to bring about the outcome against her preferences or interests were out of the ordinary.

succeeded in killing Victim$_{IN}$ *and* any one condition in this set had been missing.[28]

The fact that Intentional in (S1) is *not* metaphysically lucky with respect to Victim$_{IN}$'s death goes a long way toward defusing the misleading claim that – as compared with Failed – she is *just lucky* that her bullet hit and killed her intended victim and therefore that she should not be punished more harshly than Failed. Still, while this point helps to curb the excesses of the equivalence theorist's approach, more work still needs to be done. In particular, it still remains to be seen why the very fact that Victim$_{IN}$'s death is caused at least in part by metaphysical luck (whether or not it is on Intentional's side) is not *itself* sufficient to render Intentional blameless for it.

V. A FAILED ARGUMENT AGAINST (7)

Let us return to the Equal Punishment Argument and intuitions (1) and (2). Again, the problem is that while (1) and (2) suggest that Intentional deserves a higher degree of punishment than Failed and Unintentional deserves a higher degree of punishment than Lucky, the Equal Punishment Argument suggests that Intentional should receive the same degree of punishment as Failed and that Unintentional should receive the same degree of punishment as Lucky. So it remains to be seen why we still (should) hold on to (1) and (2) in the face of the Equal Punishment Argument. In other words, it remains to be seen what, if anything, is wrong with the Equal Punishment Argument.

It may at first be tempting to argue that it is premise (7) in the Equal Punishment Argument that is false. Again, (7) states:

In any given situation, an agent has control at most over (a) her state of mind and (b) her bodily motions. She does not have control over (c) what occurs outside her body.

How might this argument against (7) go? One might argue that the problem with (7) is part (c), that it is simply not the case that an agent

---

[28] See Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281, pp. 254–258, p. 277). Moore recognizes the distinction between metaphysical luck's being involved and being metaphysically lucky. It is less clear, however, that he recognizes the distinction between having metaphysical luck on one's side and being metaphysically lucky.

does not have any control over what happens outside of her body. The reason (7c) seems at first to be true, the "(7c)-skeptic" argues, is because we all share the intuition that an agent cannot act where she is not, that she cannot impact on objects in her environment without direct contact. But, the (7c)-skeptic argues, this metaphysical assumption about an agent's being unable to exert "action at a distance" is true only in a limited sense. It is false in another sense. It is certainly true that an agent cannot impact on objects at a distance from her *with her naked body alone*. One who disputes this proposition will argue that some agents have telepathic powers. And we are hardly going out on a limb by rejecting this claim. But the metaphysical assumption above is false to the extent that an agent can operate certain tools or devices or machines that can themselves impact on objects at a distance from her body. A prime example of such an action-at-a-distance enabler is a gun. So Intentional still *does* have control over particular events outside her body – in this case, whether or not she kills $Victim_{IN}$. By pulling the trigger, she controls whether or not $Victim_{IN}$ will die.[29]

The (7c)-skeptic's argument is clever, but it cannot be right. For it simply denies what is incontrovertible. It is incontrovertible that Intentional does not have control over what happens to the bullet – and therefore to $Victim_{IN}$ – once it leaves her gun; that once the bullet leaves her gun, it is not up to her but up to external circumstances entirely out of her control whether or not the bullet will hit $Victim_{IN}$ and whether or not, if the bullet hits $Victim_{IN}$, $Victim_{IN}$ will die.[30] For example, Intentional had no more control over the fact that no

---

[29] See Barbara Herman ("Feinberg on Luck and Failed Attempts" *Arizona Law Review* 37 (1995): 143–149, p. 144, p. 147), Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281, p. 257, p. 277, n27), Gary Watson ("Closing the Gap" *Arizona Law Review* 37 (1995): 135–141, p. 139). In the same vein, one might argue that Kobe Bryant has great control – certainly much more control than you and me – over whether or not the basketball he throws goes through the hoop and that Tiger Woods has more control than you and me over whether or not the golf ball that he hits rolls into the hole.

[30] Russell Christopher ("Appendix to George Fletcher's 'What is Punishment Imposed For?'" *Journal of Contemporary Legal Issues* 5 (1994): 111–115) makes a similar point against Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281). See also Michael J. Zimmerman ("Taking Luck Seriously" *Journal of Philosophy* 99 (2002): 553–576, p. 562).

bird intercepted her bullet than Failed had over the fact that a bird *did* intercept her bullet. And the same holds with regard to all other such "external factors", all other factors that could have prevented their bullets from hitting and/or killing their targets. To the extent that both agents lacked control over whether or not any of these external factors obtained, both lacked control over what happened to the bullets once they left their guns.

VI. THE ASSUMPTION OF RISK PROPOSAL

Still, the (7c)-skeptic was not far off. She has the right intuition, just the wrong explanation of this intuition. Her intuition is that Intentional should still be punished for Victim$_{IN}$'s death even though Failed did the same exact thing and *should not* be punished for anybody's death. And this intuition led her to conclude that Intentional must therefore have had control over whether or not Victim$_{IN}$ died. But how did she get from her intuition to this conclusion? The only plausible answer is that the (7c)-skeptic simply assumed (3). She simply assumed, that is, that an agent "deserves punishment only for that over which she has control, not for that over which she has no control". But *this* was her mistake. As it turns out, (3) is false. As long as Intentional has control over whether or not she pulls the trigger, she is morally responsible for whatever *reasonably foreseeable* harm arises from this action of hers – even if she does *not* have *control* over this reasonably foreseeable harm itself.

But how can this be? How can Intentional be morally responsible for something over which she has no control? The answer to this question is actually quite simple: Intentional *assumed the risk*. She voluntarily created a situation in which it was reasonably foreseeable that her action would earn a certain moral status from its consequences and thereby make herself deserving of punishment commensurate with this moral status. Of course, this formulation of my "Assumption of Risk Proposal" is far too quick to pass muster. So I shall spend the remainder of this section unpacking it. This "unpacking" will break down into five different parts.

A.

The first part of the Assumption of Risk Proposal says this:

(10) Intentional and Unintentional voluntarily put themselves into a situation in which the risk that another would be injured or killed was significantly increased.

The notion that Intentional and Unintentional "voluntarily put themselves" into the particular situations they did suggests that they had *control* over their assumptions of risk, that it was not a matter of metaphysical luck that Intentional pulled the trigger or that Unintentional drove recklessly. If not only what the bullet did after leaving the gun but also what Intentional did with the gun itself – if both the period before shooting as well as the period after shooting – were products of metaphysical luck, then Intentional clearly could not be held responsible for *anything*, not even the shooting itself.[31]

---

[31] Notice, this conclusion has potentially serious philosophical consequences. For it could be used in a rather ingenious way to show that nobody has control over anything. Again, the argument here is that Intentional does not control what happens to her intended target because (a) this is up to the bullet and (b) what happens to the bullet once it leaves her gun is itself up to factors outside her control. Well, one might argue, the same is true not merely of the outcome of her act of shooting but also of her act of shooting itself. For her act of shooting too is ultimately the product of factors just as outside her control as the outcome of her action – what Thomas Nagel ("Moral Luck" *Aristotelian Society*, Supp. Vol. 50 (1976): 137–152. Reprinted in Nagel (1979), 24–38, and Watson (1982), 174–186) refers to as "constitutive" luck, "luck in one's circumstances", and the "luck in how one is determined by antecedent circumstances" (what Daniel Statman ed. (*Moral Luck* (Albany: State University of New York Press, 1993)) calls "causal luck"). See Antony Duff ("Auctions, Lotteries, and the Punishment of Attempts" *Law and Philosophy* 9 (1990): 1–37), Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281), and Thomas Nagel ("Moral Luck" *Aristotelian Society*, Supp. Vol. 50 (1976): 137–152. Reprinted in Nagel (1979), 24–38, and Watson (1982), 174–186) for different versions of this argument – even though none of these three philosophers clearly embraces it in the end. The challenge to the equivalence theorist, then, is to show us why outcome luck should be singled out from these other kinds of luck, why people should not be held responsible for the results of outcome luck but should still be held responsible for the results of constitutive, circumstantial, and antecedent luck. Without this demonstration, singling out outcome luck

Since (10) equally applies to Failed and Lucky, and since we tend to think that

(11) creating these kinds of dangerous situations itself deserves some punishment[32]

it follows that, thus far, Intentional and Failed on the one hand and Unintentional and Lucky on the other deserve the same level of punishment. For each pair created exactly the same risk through identical actions, with identical intentions, and under virtually identical circumstances. So if (10) and (11) were all there were to the Assumption of Risk Proposal, the equivalence theory would go through.

B.

But, alas, there is more. The second part of the Assumption of Risk Proposal suggests:

(12) Intentional and Unintentional deserve additional punishment because the risk of injury or death that they created was actually realized.

Again, as (10) and (11) suggest, insofar as Intentional and Unintentional "crossed the line" and deliberately posed a substantial risk to

---

Footnote 31 (continued)

from the other kinds of luck seems to be perfectly arbitrary. See Leo Katz ("Why the Successful Assassin Is More Wicked than the Unsuccessful One" *California Law Review* 88 (2000): 791–812, pp. 797–798), Kimberly D. Kessler ("The Role of Luck in the Criminal Law" *University of Pennsylvania Law Review* 142 (1994): 2183–2237, pp. 2190–2191), Michael S. Moore ("The Independent Moral Significance of Wrongdoing" *Journal of Contemporary Legal Issues* (1994): 237–281), Ken O'Day ("Some Thoughts on Joel Feinberg's Modest Proposal: Is It Really Such a Modest Proposal After All?" *Arizona Law Review* 37 (1995): 243–249, pp. 244 ff.), Michael J. Zimmerman ("Taking Luck Seriously" *Journal of Philosophy* 99 (2002): 553–576, esp. pp. 570–571).

[32] The words *these kinds of* must be emphasized. An agent should not be punished merely for creating risk to another. For her creation of this risk may be justified. Rather, an agent should be punished only for creating an *unjustified* risk. One might respond that the risk must also be "substantial" (see Model Penal Code § 2.02(2)(c)). But Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30) argues for the controversial conclusion that creation of unjustified risk, whether substantial or not, should be sufficient.

another person's well-being, they deserve a certain amount of pun-
ishment – punishment for the *attempt* to kill another in Intentional's
case, punishment for recklessly endangering others in Unintentional's
case.[33] But (12) goes further. (12) suggests that whether or not Inten-
tional or Unintentional deserves any *more* punishment on top of this
should be determined entirely by whether or not their actions produced
any harm. If there was no harm, then no more punishment should be
added; if there was harm, then more punishment should be added.[34]

While equivalence theorists will have little problem with (10) or
(11), they will have a very big problem with (12). Why, they will ask,
should Intentional or Unintentional be punished more than Failed or
Lucky respectively if they created exactly the same risk as the latter?
Why should Intentional and Unintentional be punished not merely
for risk-creation but also for risk-realization? How does the fact that
they created the risk make them deserving of additional punishment if
the risk is realized? Again, it seems fair enough to punish them for
risk-creation only because they had control over it. And they had
control over it because they had control over the actions by which the
risks were created. But the same cannot be said for risk-realization.
Again, once they created the risk, they had no control over whether
or not this risk was realized. So it is unfair to punish them any more if
it was.

---

[33] This point assumes that exposing somebody to the risk of injury or
death is itself a harm if we also assume that punishment requires harm. But I
suggested in endnote 13 above that punishment may *not* require harm. For
brief discussions of this issue, see Larry Alexander ("Crime and Culpability"
The Journal of Contemporary Legal Issues 5 (1994): 1–30, pp. 17–18), Claire
Finkelstein ("Threats and Preemptive Practices" *Legal Theory* 5 (1999):
311–338, pp. 329–330), Paul H. Robinson and John M. Darley (Justice,
Liability, and Blame: Community Views and the Criminal Law (Boulder,
Colorado and Oxford: West View press, 1995): p. 33).

[34] Barbara Herman ("Feinberg on Luck and Failed Attempts" *Arizona
Law Review* 37 (1995): 143–149, pp. 148–149) and Michael S. Moore ("The
Independent Moral Significance of Wrongdoing" *Journal of Contemporary
Legal Issues* (1994): 237–281, p. 238) make a similar point. See also Arthur
Ripstein (*Equality, Responsibility, and the Law* (Cambridge: Cambridge
University press, 1999): Ch.7). Larry Alexander ("Crime and Culpability"
The Journal of Contemporary Legal Issues 5 (1994): 1–30, p. 22) and Richard
Parker ("Blame, Punishment, and the Role of Result" *American Philo-
sophical Quarterly* 21 (1984): 269–276, p. 271, p. 273) argue against this
position.

C.

The third part of the Assumption of Risk Proposal constitutes the beginning of my answer to the equivalence theorist's objection here. I propose an analogy – the gambling analogy.[35] The concept of assumption of risk justifies (1) and (2) for largely the same kind of reason that it justifies the way in which the casino treats a gambler.

The first half of the gambling analogy involves two gamblers, Winner and Loser. Suppose that Winner and Loser decide to gamble on the roulette wheel. Both put down the same exact amount of money on different numbers. After the wheel is spun, Loser ends up losing $100 and Winner winning $200. Now, it would not be at all convincing to argue that this difference in metaphysical luck should not translate into a difference in money lost or gained. For this is the "deal" that the two gamblers made with the casino when they put their money down on the roulette wheel. Gambling conventions dictate that when a person puts her money down on the roulette wheel, she is agreeing to suffer whatever monetary fate the roulette wheel randomly dictates. So by putting their money down on the roulette wheel, Winner and Loser made a sort of irreversible agreement with the casino that they will turn this money over to the casino if the ball does not bounce their way. And this agreement justified whatever monetary fate they ultimately suffered. By putting their money at risk in this way, Winner and Loser created a situation in which they deserved whatever consequences the roulette wheel delivered, whether it was risk-realization (money loss) or not (money gain).

---

[35] David Lewis ("The Punishment that Leaves Something to Chance" *Philosophy and Public Affairs* 18 (1989): 53–67) also uses gambling to make a slightly different point. He argues that if an agent gambles with another's life, then her punishment should be determined not by the outcome of this gamble but rather by the outcome of another gamble – her drawing straws that have been proportioned to the risk that she posed to this other person. So if both Intentional and Failed posed an 80% risk of death to Victim$_{IN}$ and Survivor respectively, then 80% of the straws from which Intentional and Failed must draw should correspond with the maximum possible punishment. Notice, this still leaves a 20% chance that Intentional will be punished with a less-than-maximum sentence and an 80% chance that Failed will be punished with a maximum sentence. Antony Duff ("Auctions, Lotteries, and the Punishment of Attempts" *Law and Philosophy* 9 (1990): 1–37, pp. 17–30) offers a critical discussion of Lewis's proposal.

When Intentional shot at Victim$_{IN}$ and Unintentional drove recklessly into Victim$_{UN}$, both agents spun a proverbial roulette wheel that had a good chance of landing on another person's being severely injured and/or actually dying. But the gambling analogy is not yet complete. The second half of the gambling analogy requires us to show that, just like Winner and Loser, their spinning of this proverbial roulette wheel involved some sort of *deal* as well.

Take Intentional first. When she pointed a gun in Victim$_{IN}$'s direction and pulled the trigger, Intentional knew (a) that she was significantly increasing the risk of hurting or killing Victim$_{IN}$ by shooting at her, (b) that it is morally wrong to subject Victim$_{IN}$ to these kinds of risks, and therefore (c) that she can be reasonably expected to have refrained from pulling the trigger.[36] By acting in the face of this knowledge and still pulling the trigger, Intentional made a sort of irreversible "moral deal" – a deal with "the gods" or, even better, the "casino of *morality*". The deal was that if she failed to kill Victim$_{IN}$, then the moral status of her action (pointing and shooting at Victim$_{IN}$) would be "only" morally bad, the moral status commensurate with attempting but failing to kill, and that if she actually succeeded in killing Victim$_{IN}$, then the moral status of the very same action would be much worse – "morally *egregious*", the moral status commensurate with attempting to kill *and* succeeding in killing. So just like Loser, who made the deal with the casino to let metaphysical luck decide whether her particular bet was money-enhancing or money-reducing, Intentional made the deal with the "casino of morality" to let metaphysical luck decide whether her action was "merely" morally bad or morally egregious.[37]

---

[36] Things, of course, would get much trickier if one of these three propositions – (a), (b), or (c) – was not the case. But it is unnecessary for me to deal with these complications here. For the task now is only to see whether or not the Assumption of Risk Proposal works for the pristine, paradigmatic cases (in which (a), (b), and (c) are all satisfied). Only if it does should we then, in another paper, investigate whether or not it works for the messier, non-paradigmatic cases.

[37] Jonathan Bennett ("Shooting, Killing and Dying" *Canadian Journal of Philosophy* 2 (1973): 315–323) argues that the consequences of an action, even if temporally distant from the action itself, can help to determine the *ontological* status of the action. My point here is that the harm produced by an action, even though temporally distinct from the action, can help to determine the action's *moral* status.

I shall defend this third part further in sub-section VI E. But one might argue that even if the third part works for Intentional, it does not work for Unintentional. Again, the third part suggests that when Intentional points her gun in Victim$_{IN}$'s direction and pulls the trigger, she is making a deal with the casino of morality. She is agreeing to let the moral status of her action be determined in part by the harmful consequences of her action. But the same cannot be said of Unintentional. As a matter of psychological fact, since she does not intentionally kill Victim$_{UN}$, she does *not* so agree. So in Unintentional's case, there is no deal (no contract, no agreement) between her and the casino of morality in the first place, in which case the Assumption of Risk Proposal is inapplicable.

In response to this objection, I maintain that Unintentional *did* make a deal with the casino of morality after all. For she can be reasonably expected (a) to have known that she was creating a significant risk of injury or death not only to herself but also to others by taking her eyes off the road while she was driving and (b) to have acted on this knowledge by driving more carefully. So whether or not she consciously agreed to the deal (according to which the moral status of her action would be determined in part by its consequences) does not matter any more than it matters whether or not the gambler who can be reasonably expected to know the conventions of gambling agrees to the deal (according to which she will lose her money if the roulette wheel does not go her way) when she puts her money down. If the gambler ended up losing and then complained that she did not realize that her action had put her money at risk, her plea would fall on deaf ears. For the conventions of gambling are so widely known that, even on the off chance that she is not simply lying about her ignorance, she can be reasonably expected not to have been so ignorant. So the casino is still morally entitled to take her money. Likewise, when Unintentional drove her car recklessly, she was putting her "moral money" down on the "moral roulette wheel" and thereby subjecting it to the terms set by the norms of morality, the terms according to which the moral status of her action would be determined in part by the action's harmful consequences. And, once again, she was subject to these terms whether or not she agreed to them precisely because, ex

hypothesi, she can be reasonably be expected to have known that they were in effect.[38]

D.

The fourth part of the Assumption of Risk Proposal is simply "cleaning up". While intentional killing is morally egregious, attempting to kill but failing is "only" morally bad. Likewise, while carelessly driving without hitting anybody is "stupid", "negligent", "not so bad", carelessly driving and thereby killing another is "reckless", "reprehensible", "appalling". (Consider, for example, what we tend to say about drunken drivers who end up killing others as opposed to drunken drivers who do not.) Given these moral distinctions and the fact that the kind of treatment one deserves generally correlates with the moral status of her action,[39] Intentional deserves harsher punishment than Failed and Unintentional deserves harsher punishment than Lucky.

One might offer the following reductio against the latter point about Unintentional.[40] Assume I am correct that (a) unintended but reasonably foreseeable harm may be factored into an agent's desert and therefore punishment. Then (b) by parity of reasoning, (c) unintentional but reasonably foreseeable *benefit* should also be factored into an agent's desert and therefore treatment as well. But (c) is false. Therefore (a) is false. My response, however, is to reject not (a) but (b). There is an asymmetry between how we should treat unintended but reasonably foreseeable harm and how we should treat unintended but reasonably foreseeable benefit. Only the former, not the latter, should be factored into an agent's desert and treatment because society has a greater interest in minimizing actions

---

[38] In sub-section VI E, I shall lend brief support to "ethical intuitionism", the theory that the terms of morality – moral beliefs or judgments – are justified by our moral intuitions.

[39] At least according to retributivist accounts and the more respectable consequentialist accounts. See the fourth consideration in section III above. Herbert Fingarette ("Punishment and Suffering" *Proceedings and Addresses of The American Philosophical Association* 50 (1977): 499–525, pp. 511–512), however, is skeptical about this point.

[40] Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30, pp. 12–14) inspired me to think of this objection.

significantly increasing the probability of harm than in maximizing actions significantly increasing the probability of benefit.

E.

In the third part of the Assumption of Risk Proposal, I assumed that Intentional and Unintentional entered into a deal with the casino of morality according to which the harm produced by their actions will help to determine their actions' moral status. Now, to say that one enters into a deal with morality according to which something is the case is just to say that morality *requires* this something to be the case and therefore that this something *should* be the case. So if Intentional pulls the trigger and thereby "seals the deal" with morality, then the harm produced by Intentional's action *should* help to determine her action's moral status. But if I left it there, I might very well be charged with begging the question. For this assumption is mostly what is at stake in the debate between equivalence and non-equivalence theorists. So the purpose of this last part – the fifth part of the Assumption of Risk Proposal – is to avoid this charge, to support this assumption, to show why morality would require Intentional or Unintentional to make *this* deal in the first place. I shall offer four reasons – one meta-ethical, two ethical, and one pragmatic.

The meta-ethical reason why morality would require Intentional and Unintentional to enter into such a deal is the theory of "ethical intuitionism". Ethical intuitionism says that our moral intuitions directly justify the moral beliefs or judgments that arise from them.[41] Since this theory is plausible, one good reason to believe that the harm produced by Intentional's pulling the trigger and by Uninten-

---

[41] For different accounts of ethical intuitionism, see Robert L Frazier ("Intuitionism in Ethics" In Edward Craig ed. (*Routledge Encyclopedia of Philosophy* (London: Routledge, 1998) 853–856, Peter Railton ("Analytic Ethics", In Craig (1998) 220–223, pp. 220–221), John Skorupski ("Ethics" In Nicholas Bunnin and Eric Tsui-James (eds.) (*The Blackwell Companion to Philosophy* (Oxford: Blackwell Publishers Ltd., 1996)), 202–230, pp. 219–221). See also the collection of essays in Philip Stratton-Lake ed. (*Ethical Intuitionism: Re-evaluations* (Oxford: Oxford University Press, 2002)). Paul H. Robinson and John M. Darley (*Justice, Liability, and Blame: Community Views and the Criminal Law* (Boulder, Colorado and Oxford: Westview Press, 1995), p. 6) offer some considerations for and some considerations against this theory, though they do not refer to it as such.

tional's driving recklessly should help to determine these actions' moral status is simply that this proposition is supported by our moral intuitions – i.e., (1) and (2).

We have already come across the first ethical reason why morality would require Intentional and Unintentional to enter into a deal according to which the harm produced by their actions would help to determine their actions' moral status. It is Duff's argument, which I summarized in section III. Again, Duff defends intuitions (1) and (2) by arguing that (a) the amount of harm that an agent causes *does* matter; (b) (a) *should* be communicated to the agent and to the wider community; and (c) the best, if not the only, way to accomplish (b) is by factoring the amount of harm that the agent causes into her punishment.[42] The only problems that I attributed to this argument – namely, the "message" problem and the "fairness" problem – do not vitiate it for our purposes. Regarding the message problem, I suggested above that punishment sends a message to society and the perpetrator not merely about blameworthiness but also about wrongdoing. And regarding the fairness argument, I shall address it below in Section VIII (in my Replies to Objections 4 and 5). My argument will be that fairness and desert are largely independent of one another; that even if it were unfair, in certain circumstances, to punish Intentional and Unintentional additionally for the harm that they produced, they would still fully deserve this additional punishment.

The second ethical reason attempts to hoist the equivalence theorist on her own petard. We have already seen in sub-section VI A that the equivalence theorist believes that Intentional should be punished for pulling the trigger. But why should she be *punished* for this action rather than *rewarded* or simply *left alone*? Because there is

---

[42] Incidentally, Duff's argument provides an effective reply to Alexander's challenge to non-equivalence theorists to explain why punishment should be affected by one kind of metaphysical luck (i.e., the harm produced by one's action) but not another kind of metaphysical luck (e.g., drawing lots). See Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30, p. 28). The difference is (a). Harm matters. It has intrinsic significance. Which straw is picked does not. I make a similar point in section VIII in my Reply to Objection 2.

a law against pulling triggers.[43] But why is there a law against pulling triggers? Because such actions tend to hurt or kill others. So it is the *reasonably foreseeable* consequences of this kind of action that help to determine what moral status this particular action has and therefore what kinds of laws there are with respect to this kind of action and therefore what kind of treatment it deserves from the state – punishment or reward or nothing at all.[44] It is unclear, then, why the equivalence theorist would concede this point but not (12); why she thinks that it is perfectly reasonable to allow only the *reasonably foreseeable*, not the *actual*, consequences of an action to determine its moral status – *especially* when the actual *match* the reasonably foreseeable in the first place. If reasonably foreseeable consequences can help to determine the moral status of Intentional's pulling the trigger, then certainly the fulfillment of this reasonable foreseeability can do the same.

Finally, the pragmatic reason why morality would require Intentional and Unintentional to agree to let the results of their actions help to determine their actions' moral status depends on an extension of the gambling analogy. We may think of morality *as if* it is an agent with intentions just like the real casino. That is, just as the real casino enters into a deal with the gambler in the hopes of acquiring the right to the gambler's money, we may think of morality *as if* it is an agent entering into a deal with Intentional and Unintentional to promote *its own* goals. What might these goals be? Well, given that it is *morality* that is the (as if) agent here, its primary goal must ultimately be to promote good behavior, doing the right thing, action that is in

---

[43] Herbert Fingarette ("Punishment and Suffering" *Proceedings and Addresses of The American Philosophical Association* 50 (1977): 499–525, pp. 508 ff.) argues that the very concept of legal obligation (what Fingarette sometimes refers to as "legal requirement" and the "power of law") entails retributive punishment for voluntary violations of this obligation. For a different justification of retributivism, a justification based on our natural retributive sentiments, see Jonathan Jacobs ("Luck and Retribution" *Philosophy* 74 (1999): 535–555).

[44] See David McCarthy ("Harming and Allowing Harm" *Ethics* 110 (2000): 749–779, pp. 749–750). Another way to put the suggestion here is that the action is wrong, and therefore should be punished, because it wrongfully significantly increases the risk of harm. I leave it an open question, however, whether or not in addition to being wrongful, increasing the risk of harm is *itself* a harm. See endnote 33.

accord with its own principles. Morality is most happy when a person does what it prescribes. So this is precisely the goal at which it aims. And because morality is not stupid, it may very well reason that it will do a better job of achieving this goal, of maximizing good behavior and minimizing bad behavior, by forcing a person who wishes to *violate* its principles to enter into a deal with the toughest possible terms. Now, certainly a deal which says that the harm produced by a would-be offender's action will help to determine her action's moral status is tougher than a deal which says that her action's moral status will be fixed at the time of the action. The former deal is tougher than the latter because the former, unlike the latter, forces the would-be offender to take a risk. It says that if she goes ahead and violates morality by pulling the trigger, then she is not merely scoring "immoral points" for this action itself but is also risking the addition of further immoral points should the action lead to any harmful consequences. Naturally, some would-be violators might be discouraged by this risk (of scoring further immoral points) alone from carrying through with the violation. Moreover, those who are not discouraged by this risk alone might still be effectively discouraged by this risk *in conjunction with* the risk of additional punishment – as I advocate in the fourth part of the Assumption of Risk Proposal above. (I shall further discuss similar deterrence-related points in sub-section VII E.) Finally, morality might still be satisfied that those who are not discouraged by these two risks may very well end up being punished more harshly than they would have been had they violated its precepts *without* entering into any such deal.

F.

The Assumption of Risk Proposal has implications that extend beyond Intentional and Unintentional. It helps to explain and justify not merely punishing successful attempts more than failed attempts and harmful carelessness more than harmless carelessness but, more generally, behavior leading to harm that was a reasonably foreseeable outcome of such behavior more than the very same behavior leading to less or no harm. Indeed, this principle is hardly a stranger to legal doctrine. In both criminal and tort law, the general rule is that a defendant may be found liable for the harm caused by her wrongdoing if and only if this harm was a reasonably foreseeable consequence of

her wrongdoing.[45] So the Assumption of Risk Proposal may be used to explain and justify, for example, punishing an accomplice more for the reasonably foreseeable harm caused by the perpetrator than she would have been punished had no such additional harm resulted – even if the accomplice in question was not present at the scene of the crime and did not want, intend, or expect the perpetrator to cause additional harm.[46]

VII. AN ARGUMENT AGAINST (3)

In the last section, I tried to satisfy the burden that the equivalence theorist poses to the non-equivalence theorist. I tried to justify the proposition that an agent's moral desert may be determined in part by the harm produced by her action even though she may not have control over whether or not her action produces any harm. In this section, however, I wish to shift the burden. I wish to force the *equivalence theorist* to justify her own position that

(3) an agent deserves punishment only for that over which she has control, not for that over which she has no control.

I shall argue that there is no good justification for (3) and therefore that, whether or not she realizes (and likes) it, the equivalence theorist has no good reason to reject the Assumption of Risk Proposal.

A.

Another way to state (3) is this: in order for an agent to be deserving of punishment for any given thing X, it must at least be the case that she had control over X. But *why*? Just *why* should we accept (3)? When it comes to punishing Intentional for X, why should it matter

---

[45] As usual, there are exceptions to this general rule. In criminal law, intent trumps foreseeability; a defendant may be found liable for unforeseeable harm as long as she intended it. And in tort law, the so-called "eggshell skull doctrine" says that a tortfeasor may be liable for harm to her victim even if this harm resulted in part from the victim's unforeseeable fragility.

[46] See, e.g., The People v. Luparello, 187 Cal. App. 3d 410, 231 Cal. Rptr. 832 (Cal. Ct. App. 1986).

whether or not she has control over X?[47] What are the justifications for regarding control over X as a necessary condition for punishing for X?. There are two kinds of answers that the equivalence theorist (or anybody else who subscribes to (3)) may give to these questions. First, she may argue that:

(13) there just *is* no deeper reason in virtue of which control is necessary for punishment.[48] (3) is undeserved a brute moral fact.

Second, she may argue that:

(14) control is a necessary condition of punishment by transitivity. Punishment requires *something else* that itself requires control.

Clearly, if she offers (14), we must find out what this "something else" might be. I can think of only four plausible candidates. Regarding Intentional and the punishment that she should suffer for pulling the trigger, they are:

(14a) she could have avoided or refrained from pulling the trigger;
(14b) she can be reasonably expected to have avoided or refrained from pulling the trigger;
(14c) she was responsible for pulling the trigger;

and/or

(14d) she can be deterred from pulling future triggers.

The task now is to see whether either (13) is true or any of (14a) through (14d) both requires control and is itself required by punishment. I shall argue that neither is the case and therefore that the equivalence theorist has no good reason to subscribe to (3).

B.

(13) suggests that there just *is* no justification for (3), no deeper reason in virtue of which (3) is true. Rather, the notion that agents can and should be punished for the wrongful actions over which they have control constitutes nothing more – and nothing less – than a

---

[47] I am confining my discussion in this section to Intentional and leaving out Unintentional for the sake of simplicity.

[48] That is, necessary for *deserved* punishment. Likewise, in what follows, I am concerned only with deserved punishment, not undeserved punishment.

basic moral truth. The problem with this proposal, however, is that every moral truth derives ultimately either from other even deeper moral truths or from a theory of value (i.e., a theory of what has value and why).[49] If the equivalence theorist's proposal here is that (3) does not derive from a deeper moral truth, this assumption still cannot be the end of the story. For we are still entitled to know what value it helps to preserve or strengthen. And since I cannot imagine any other value that (3) might help to preserve or strengthen than (14c) and/or (14d), (3)'s fate must ultimately depend on the fate of these two propositions (which I discuss below).

C.

Possibly (14a) and certainly (14b) are both required by punishment. But neither of them requires control. For there are circumstances over which one does not have control that one can still avoid and can still be reasonably expected to avoid. The tragic consequence of Intentional's pulling the trigger provides a perfect example. It is quite clear that Intentional could very well have avoided and can be reasonably expected to have avoided producing the consequence of Victim$_{IN}$'s dying. Yet it is agreed on both sides that, once the bullet left her gun, she did not have control over this consequence (see section V).[50]

It may seem counter-intuitive to suggest that an agent could have avoided and can be reasonably expected to have avoided some consequence C without having control over C. But it really is not as counter-intuitive as it may first appear. For there are two other conditions that must be satisfied, and these two conditions help to make the absence of control over C much less troublesome. The two other conditions are that the agent had control over some *causal precedent* of C (in this case, pulling the trigger) *and* that C is a *reasonably foreseeable* consequence of this causal precedent. In other

---

[49] This statement is consistent with my endorsement of ethical intuitionism in sub-section VI E. Moral truths derive ultimately from basic moral truths, basic moral truths from theories of value, and theories of value from our moral intuitions.

[50] Another example that helps to prove the same point is the "self-made" alcoholic. While she may not have control over her present act of drinking, she is still responsible for it. For we think that she could have avoided, and can be reasonably expected to have avoided, becoming an alcoholic in the first place.

words, the agent *could* have avoided and can be reasonably expected to have avoided C not only if she directly produced C but even if she "only" *in*directly produced C – i.e., directly produced a causal precedent of C from which C was reasonably foreseeable.

Notice, this proposal is inconsistent with the equivalence theory. For this proposal is inconsistent with (3), the proposition that punishment for X requires control over the very same entity – X. Indeed, if the equivalence theorist were now to "dilute" this assumption by allowing punishment for X even if the agent has control only over a causal precedent of X and not over X itself, then it would no longer be clear why she thinks that Intentional should not be punished any more for killing Victim$_{IN}$ than for attempting to kill Victim$_{IN}$ in the first place.

D.

(14c) suggests that control is necessary for punishment because responsibility requires control, and the (or a) purpose of punishment is to honor or celebrate responsibility. Clearly, proponents of (14c) have in mind *moral* responsibility here. But it is difficult to see what more moral responsibility involves here than *causal* responsibility plus conditions (14a) and (14b). We have already seen that (14a) and (14b) fail as explanations of (3) because neither requires control over C, only control over some causal precedent of C. So we are left with causal responsibility. But it turns out that causal responsibility runs into the same problem. An agent can be causally responsible for C even she did not directly bring about C. Intentional is clearly a key contributing cause of Victim$_{IN}$'s death. If she had not pulled the trigger, Victim$_{IN}$ would not have died at that time in that way.[51] Yet Intentional cannot be said to have control over Victim$_{IN}$'s fate because she had no control over the bullet once it left the gun. So just as with (14a) and (14b), we seem to run into the counter-intuitive conclusion that moral responsibility for C does not require control over C. But also like (14a) and (14b), the counter-intuitiveness of this

---

[51] Of course, this statement might be false if somebody other than Intentional had shot an equally fatal bullet at Victim$_{IN}$ at the same time. But we need not worry about cases of causal overdetermination yet. As I mentioned in endnote 36, the task now is only to see whether or not the Assumption of Risk Proposal works for the paradigmatic cases, not whether or not it works for such non-paradigmatic cases as causal overdetermination.

suggestion is mitigated by the fact that moral responsibility for C *still* requires control – just not control over C. Again, like (14a) and (14b), it entails control over some causal precedent of C from which it was reasonably foreseeable that C would result.

E.

Finally, (14d) suggests that punishment for the purpose of deterrence requires control because an agent cannot be deterred unless she can change her behavior, and an agent cannot change her behavior unless she has some minimal degree of control over it. But, first, it is not clear that punishment is justified only if it deters or is likely to deter. Even if a given punishment were conclusively proven to have no deterrent effect, it might arguably still be morally legitimate as long as the defendant *deserved* it. Second, even if we concede the highly controversial premise that punishment is not justified unless it deters or is likely to deter, deterrence does not seem to require control.[52] This conclusion applies not only to control over consequences but also to control over the attempt itself.

Consider first control over the attempt. Even if Intentional lacked control over her attempt to kill Victim$_{IN}$, it might still serve the purposes of both general and specific deterrence to punish her for this attempt. Regarding general deterrence, even if Intentional lacked control over her attempt, many other would-be attempters may not lack such control. And regarding specific deterrence, just because (ex hypothesi) Intentional lacked control over this particular attempt does not necessarily mean that she will lack control over all future attempts. So even if she lacked control over this particular attempt, punishing her for it may still work to deter her from making similar kinds of attempts in the future.

Likewise with control over consequences. Even if we concede that Intentional has no control over whether or not Victim$_{IN}$ dies, punishing Intentional not just for attempting to kill Victim$_{IN}$ but also for Victim$_{IN}$'s death would arguably deter her moreso from attempting to kill others in the future than would punishing Intentional just for *attempting* to kill Victim$_{IN}$. More generally, a system that punishes successes in addition to attempts will arguably be more effective in

---

[52] Contrary to Larry Alexander ("Crime and Culpability" *The Journal of Contemporary Legal Issues* 5 (1994): 1–30, p. 7).

298 KEN LEVY

deterring individuals from making such attempts in the first place than a system that punishes only the attempts themselves. For if only attempts were punished, then the only threat success would pose to an attemptor is the evidence of attempt that successes tend to leave behind. But if success were also punished, then any given attemptor would face not just one but *two* threats – *both* leaving behind possible evidence of the attempt *and* an increase in overall punishment (One might respond to all this that punishing success in addition to attempts will not deter attempts. Rather, it will simply encourage attempts over success. But this suggestion is absurd. One cannot be encouraged to attempt and discouraged from success at the same time. To attempt is just to attempt to succeed.).[53]

F.

Given the Assumption of Risk Proposal and my arguments in this section, the burden is now on the equivalence theorist to show why C's being an actual and reasonably foreseeable consequence of my action is *not* sufficient for C's being factored into the moral status of my action. What does direct control over C have that "indirect control" – i.e., control over a causal precedent of C from which C was a reasonably foreseeable consequence – does not? Why should we think that I deserve control only for things over which I have direct control and not for things over which I have only indirect control? It is my contention that the equivalence theorist simply cannot provide a satisfactory answer to these questions.

VIII. OBJECTIONS AND REPLIES

I can think of several different objections that the equivalence theorist might raise against the Assumption of Risk Proposal. I shall explicate and address them here.

Objection 1: Since the harm produced by an action occurs after the action, the notion that the former may affect the moral status of the latter requires backward causation. But backward causation is metaphysically impossible. Therefore the harm produced by an action may not affect the moral status of the action.

---

[53] See Gerald Dworkin and David Blumenfeld ("Punishment for Intentions" *Mind* 75 (1966): 396–404, p. 398).

Reply: It is by no means clear that backward causation is metaphysically impossible. Even if it is difficult to stomach in the physical realm, it is quite easy to accept in the conceptual realm. Indeed, we have no problem with the notion that the roulette ball's final position on the roulette wheel retroactively determines the profit status of the gambler's earlier bet. Likewise, then, we should have no more problem with the notion that the harm produced by an action retroactively helps to determine the moral status of the action.

Objection 2: The gambling analogy is not strong enough to support the Assumption of Risk Proposal. For it is outweighed by a fundamental *dis*analogy between gambling and the criminal justice system. Quite simply, while gambling is just a *game*, the criminal justice system is anything but a game. Gambling's ultimate purpose is to provide its participants with *fun, relaxation, enjoyment*, and perhaps the opportunity for *easy money*. Given such morally frivolous purposes, the consequences for any given individual of participation in this game may be considered equally morally frivolous as well. And the more morally frivolous the consequences, the more morally acceptable it is to allow luck to play a role in determining these consequences. Conversely, then, the more morally "momentous" the consequences, the *less* morally acceptable it is to allow luck to play a role in determining these consequences. And how we characterize the moral status of a given action as well as how much we punish that individual as a result of the moral status we assign to her action certainly qualifies as a morally momentous consequence. It is not a game but rather a very serious business. So while it is appropriate to speak about gambling with one's money, it is entirely inappropriate to speak about gambling with the moral status of one's action.

Reply: To be sure, it would not be justified to let *any* kind of metaphysical luck – e.g., the bounce of a roulette ball – help to determine an agent's degree of punishment. But it *is* justified to let *morally related* metaphysical luck – i.e., metaphysical luck that affects an agent's moral desert – help to determine this degree. And the metaphysical luck that contributed to Victim$_{IN}$'s and Victim$_{UN}$'s deaths certainly affects Intentional's and Unintentional's moral desert. For each voluntarily brought about the situation that helped to create this life-endangering metaphysical luck in the first place.

Objection 3: Assume that Intentional and Unintentional live in a society that is governed by an "Equal Punishment System", not a "Differential Punishment System", and that they know this fact. The criminal code clearly states that people like Failed and Lucky will be punished just as harshly as people like Intentional and Unintentional respectively, and both Intentional and Unintentional know that the criminal code contains these provisions. In this situation, which is perfectly possible, the Assumption of Risk Proposal – in particular, the third part – must be wrong. For in this situation, neither Intentional nor Unintentional agrees to let the harm produced by their actions affect their moral desert. Instead, given that they both know that they are in an Equal Punishment System, both make the very *opposite* deal – the deal that whatever harm is produced by their actions will *not* affect their moral desert.

Reply: The third part of the Assumption of Risk Proposal does *not* suggest that Intentional's and Unintentional's moral desert is determined by any deal that they make with *society*. Rather, it suggests that their moral desert is determined only by the deal that they make with the *casino of morality*. And the deals that morality arranges with would-be offenders are independent of whatever punishment system that society has put in place. Their terms remain the same, whether the system in place is Differential Punishment or Equal Punishment. So even in an Equal Punishment System, Intentional's and Unintentional's moral desert will *still* be affected by the harm produced by their actions.

Objection 4: Their moral desert perhaps. But *not* their punishment. Given that Intentional and Unintentional are, ex hypothesi, living in an Equal Punishment System, it is now entirely *unfair* to punish Intentional more than Failed or Unintentional more than Lucky. For neither had fair notice of the fact that their killing others might bring them greater punishment than their merely endangering others. On the contrary, both were given notice that their punishment would *not* be affected by the harmful consequences of their actions; that if they are convicted of first-degree homicide or reckless homicide, they will be punished just as harshly as they would have been had they been convicted of attempted homicide or reckless endangerment respectively. So the Assumption of Risk Proposal works only in a Differential Punishment System, *not* in an Equal Punishment System. It would lead to fair results only in the former, not in the latter. It

therefore begs the question against the equivalence theory. It fails to tell us why, if we happened to find ourselves in an Equal Punishment System, we should still move to a Differential Punishment System.

Reply: I concede the equivalence theorist's point that it would be unfair to punish Intentional and Unintentional more harshly than Failed and Lucky respectively in a well-publicized Equal Punishment System. But here we need to return to the distinction that I drew (and downplayed) in Section III above, the distinction between fairness and desert. Even if it might be unfair to punish Intentional and Unintentional more than Failed and Lucky respectively, they would still be getting precisely what they *deserve*. For unlike fairness, which is essentially comparative, desert is essentially non-comparative.[54] The punishment that Intentional and Unintentional deserve is entirely independent of the punishment that Failed and Lucky deserve or receive. What Intentional and Unintentional deserve depends entirely on their actions and, if the Assumption of Risk Proposal is correct, the harm produced by their actions. It does not at all depend on anything involving Failed or Lucky – their actions, the harm produced by their actions, how much punishment they deserve, or how much punishment they receive.[55]

---

[54] I am assuming that fairness and desert are mind-independent entities. But even mind-independent entities sometimes require minds to "figure them out". So even if I am right that desert is essentially non-comparative, comparison may still be required for us to figure out what this mind-independent desert is in any given situation.

[55] Christopher J. Peters ("Foolish Consistency: On Equality, Integrity, and Justice in Stare Decisis" *Yale Law Journal* 105 (1996): 2031–2115) and Peter Westen (*Speaking of Equality* (Princeton: Princeton University Press, 1990), Ch. 9) make a parallel argument in the context of political theory. Both argue (a) that each individual deserves to be treated in a certain kind of way by the state and (b) what determines how she should be treated depends only on her own situation and the relevant moral and political considerations. Two things follow from both of these premises. First, the state should treat two individuals who are similarly situated in all morally and politically relevant respects the same. Second, contrary to popular wisdom, the reason that they should be treated the same is *not* because it would be unfair to one of them to treat the other better. Rather, the reason that they should be treated the same is because (a) and (b) together entail that each independently deserves the same kind of treatment. So to treat them differently would mean that at least one of them was not getting the kind of treatment that she deserves.

In the end, which kind of punishment system is in place, Equal Punishment or Differential, affects only the fairness or unfairness of factoring harm into an agent's punishment. It does not affect the agent's moral desert. So even if Intentional and Unintentional killed in an Equal Punishment System, they would still be getting what they deserve if, contrary to the Equal Punishment provisions and therefore unfairly, the harm produced by their actions was factored into their punishment.

Objection 5: I have not delivered what I promised in Section III above. Again, in section III, I suggested that I would "complement Duff's argument" for the conclusion that harm *should* be factored into one's punishment by showing "that it is indeed fair to factor into one's punishment this harm that matters." But it looks as though I have just conceded in my Reply to Objection 4 that it *is not* fair to factor harm into Intentional's or Unintentional's punishment.

Reply: First, when I made that promise in Section III above, I was not promising to show that factoring harm into one's punishment is and always must be perfectly fair in all respects. What I was promising to show was something much more specific – namely, that factoring harm into one's punishment may *not* be considered *unfair simply for the reason that* this harm was out of the agent's control. This promise still left perfectly open the possibility that I conceded in my Reply to Objection 4 – namely, that factoring harm into an agent's punishment may be considered unfair for *another* reason, the reason that she was not given fair notice that such factoring would or might occur. Again, if Intentional or Unintentional acted as they did in an Equal Punishment System and knew that they were in an Equal Punishment System, factoring any harm into their punishment *would* be unfair.

Second, there are different kinds of fairness. So the fact that it is unfair in one respect to punish Intentional more than Failed does not mean that it is unfair in *every other* respect as well. On the contrary, in addition to fair notice, another kind of fairness is getting what one deserves. We say that the teacher is being perfectly fair to Johnny when she gives him less candy than she gives to Sally because Sally was a good girl and Johnny a bad boy. Likewise, we say that life is unfair when we think that we deserve better than what we are getting. So if I have succeeded in showing that

Intentional and Unintentional get what they deserve when they are
punished more than Failed and Lucky respectively, whether in an
Equal Punishment System or in a Differential Punishment System,
then I have also succeeded in showing in the *desert sense* of fairness
that factoring harm into their punishment is indeed fair *even if* this
harm was not in their control.

IX. CONCLUSION

I conclude that our intuitions – (1) and (2) – remain standing. The
biggest threat to (1) and (2) – namely, the Equal Punishment Argu-
ment – rests on a mistaken assumption about the relation between
metaphysical luck and desert. In the end, Intentional and Uninten-
tional should be punished more than Failed and Lucky respectively
because they caused greater harm than Failed and Lucky. And even
though this difference in the outcome of their actions was due to
factors outside their control, differences in metaphysical luck, both
Intentional and Unintentional assumed the risk. Each agent volun-
tarily created a situation that she knew or at least should have known
would let the moral status of her action be determined by one or
another reasonably foreseeable outcome of her action. So like the
roulette-wheel gambler, who deserves whatever reasonably foresee-
able consequences the roulette wheel brings her (i.e., whether losing
her money or winning more), Intentional, Failed, Unintentional, and
Lucky deserve whatever punishment is proportional to the moral
status of their spinning this proverbial roulette wheel, where the
moral status of their spinning is itself determined in part by the harm
(if any) resulting from this spinning.[56]

225 E. 95th street
Apt. 27M
New York, NY 10128
USA
E-mail: KLevy2001@yahoo.com

---

[56] I would like to thank Russell Christopher, Ken Ehrenberg, George
Fletcher, John Gardner, Michael Siebecker, and anonymous referees at Law
and Philosophy for their helpful comments on earlier drafts of this paper.