

Unwitting Justification

PETER WESTEN*

TABLE OF CONTENTS

I.	INTRODUCTION	420
II.	THE CONDITIONS OF UNWITTING JUSTIFICATION	421
	A. <i>Jill and Earl</i>	423
	B. <i>Jurisdictions X and Y</i>	424
III.	WHY HARMFUL RESULTS MATTER	425
IV.	APPLYING PLATO’S RATIONALE FOR RESULTING HARM TO UNWITTING JUSTIFICATION	428
	A. <i>Positive Relief</i>	428
	B. <i>Lack of Regret</i>	430
	1. <i>Jill-3 and Earl-3</i>	430
	2. <i>Man on the Track</i>	432
	3. <i>Unwitting Parent in “Man on the Track”</i>	432
	C. <i>Regret</i>	433
	1. <i>Tactical Bombers</i>	434
	2. <i>Unwitting Villagers</i>	434
	3. <i>Breaching a Levy to Save a Town</i>	435
V.	STEINHOFF’S ARGUMENTS AGAINST OBJECTIVE JUSTIFICATION	436
	A. <i>Objective Justification Misconstrues the Linguistic Meaning of Justification</i>	436
	B. <i>Objective Justification Requires That Unwittingly Justified Actors Get Off the Hook Entirely</i>	437
	C. <i>Objective Justification Makes Justified Self-Defense Practically Impossible</i>	442
	1. <i>Steinhoff’s Critique Applies as Much to His Own Theory of Justification as to Objectivism</i>	443

* © 2018 Peter Westen. Frank G. Millard Professor of Law, Emeritus, University of Michigan.

2.	<i>Steinhoff's Critique is Based on a Mistaken Notion of Necessity</i>	444
D.	<i>Objective Justification Fails to Provide Action Guidance</i>	445
VI.	CONCLUSION	446

I. INTRODUCTION

We criminal law theorists are a strange lot. I doubt any of us can document an actual instance of unwitting justification, other, perhaps, than the 1850 case, *R. v. Dadson*.¹ Yet we continue to engage one another about whether mental states of knowledge or belief are elements of the defense of justification, ringing changes on arguments that George Fletcher and Paul Robinson made in debating one another forty-two years ago.²

We presumably continue the debate because, despite it having little practical import, the issue appears to call into question basic ways in which we conceptualize criminal responsibility. It is commonplace, for example, to conceptualize criminal responsibility as consisting of two elements—wrongdoing and culpability—where culpability, like *mens rea*, consists exclusively of an actor's mental capacities and states, and wrongdoing, like *actus reus*, consists of something distinct from that.³ Yet the debate over unwitting justification poses the question whether wrongdoing itself consists in part of the mental states with which a person acts. Similarly, with respect to completed offenses, it is commonplace to conceptualize wrongdoing as consisting of the objective harms and risky acts that the state deploys criminal statutes to prevent, all things considered, while culpability consists of an actor's mental capacities and states regarding those harms and risky acts.⁴ Yet unwitting justification poses the question whether wrongdoing, all things considered, also takes account of an actor's mental state prior to considering culpability. Finally, it is commonplace to regard completed offenses as the paradigms of criminal conduct and to regard criminal attempts as normatively—in addition to conceptually—derivative of them. Yet, as we shall see, unwitting justification calls that assumption into question, too.

1. *R. v. Dadson* (1850) 169 Eng. Rep. 407, 407; 2 Den. 35, 35 (“The prisoner was not justified in firing at Waters, because the fact that Waters was committing a felony was not known to the prisoner at the time.”).

2. See Paul H. Robinson, *A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability*, 23 UCLA L. REV. 266, 288–91 (1975); see also George P. Fletcher, *The Right Deed for the Wrong Reason: A Reply to Mr. Robinson*, 23 UCLA L. REV. 293, 318–21 (1975).

3. See, e.g., Heidi M. Hurd, *Justification and Excuse, Wrongdoing and Culpability*, 74 NOTRE DAME L. REV. 1551, 1559 (1999).

4. *Id.* at 1558–59.

I am among those who believe unwitting justification is a revealing window into criminal responsibility. However, the issue concerns me for another reason as well: our collective inability, as theorists, to come to consensus after more than forty years of debate makes me wonder what we can ever possibly agree upon, if we cannot agree about this. Personally, I have always found Paul Robinson's position so compelling in comparison with his opponents' position that I wonder whether our continuing disagreements mask emotions on all our parts that reason cannot penetrate.⁵ Still, I turn to reasoned argument because reason is the only tool we theorists possess.

I proceed in four parts. Part II sets out the conditions of unwitting justification and argues that (i) regardless of whether unwittingly justified offenders are guilty of completed offenses of unjustified harm, they are clearly guilty of attempting to commit such offenses, and (ii) whether they should be punished for completed crimes of harm rather than attempt depends upon *why* criminal law punishes completed offenses more severely than impossibility attempts. Part III argues that the best account of why harmful results matter is one that Plato advanced 2300 years ago in *The Laws*. Part IV applies Plato's account to three sets of cases of unwitting justification and argues that, by Plato's account, at least two of them should be treated as attempts rather than as completed crimes. Finally, Part V analyzes Uwe Steinhoff's claim that unwittingly justified offenders should be punished for completed crimes of harm rather than for crimes of attempt.

II. THE CONDITIONS OF UNWITTING JUSTIFICATION

Consider the various ways in which persons can succeed or fail to inflict harm on others. A person can inflict an *unjustified* harm on another without intending knowing, or—consciously or negligently—risking an unjustified harm, for example, a target-range shooter who hits the bullseye at a shooting range, not intending, knowing, or—consciously or negligently—risking that he is killing a child who has surreptitiously hidden himself behind the target. A person can also inflict a *justified* harm on another without intending, knowing, or—consciously or negligently—risking that he is harming anyone at all, for example, a target shooter who hits the bullseye at a shooting range, not intending, knowing, or—consciously or negligently—risking that he is killing a terrorist who has concealed himself behind the target and

5. See Robinson, *supra* note 2.

is about to explode a device that will kill everyone at the range.⁶ All such actors lack moral and legal culpability because they do not intend, know, or—consciously or negligently—risk that they are killing human beings.

Conversely, a person can *fail* to inflict an *unjustified* harm on another while nevertheless intending, believing, or—consciously or negligently—risking such unjustified harm, for example, a person who shoots at a rival out of jealousy but misses. A person can also fail to inflict *justified* harm on another while nevertheless intending, believing, or—consciously or negligently—risking an unjustified harm, for example, a person who shoots at but misses a rival while unaware that the rival is about to try, albeit unsuccessfully, to kill him. All such actors are morally culpable of impossibility attempts, and all of them could be punished for attempt under properly framed criminal laws.⁷

The issue of unwitting justification involves a third set of cases. A person, A, commits an act of *unwitting justification*, as I use the term, when the following conditions are met: (1) A harms a third person, B; (2) A intends, knows, or is reckless or negligent regarding the fact that he is harming B; (3) the harm that A inflicts upon B is proportionate and necessary to protect A or others either from imminent harm by B or from a greater evil; and (4) A does not intend or know, and is not conscious of, the likelihood that inflicting the harm is necessary to protect himself or others from such harm or evil. Condition 1 is included because, unless an actor actually harms another, the issue of his being justified in inflicting harm does not arise. Condition 2 is included because, absent condition 2, an actor is not culpable of any offense, much less a completed offense. Condition 3 is included because it specifies what I mean by *justification* for A's inflicting harm on B. And condition 4 is included because, absent condition 4, justification is not *unwitting*.

A note on terminology. I am aware that for subjectivists, inserting *justification* into the term *unwitting justification* appears to beg the question by presupposing the very norm at issue. However, I do not use the term for that purpose. I use *unwitting justification* solely for the purpose of referring to conditions 1–4, not for the purpose of deciding whether persons who act under those conditions should have a defense to completed crimes of harm. I use the term to refer to conditions that *raise* the latter question, not conditions that *resolve* it. In addition, and in the interest of economy of expression, I confine myself to actors who *intend* or *know* that they are harming third

6. See Uwe Steinhoff, *Just War Theory: Self-Defense, Necessity, and the Ethics of Armed Conflicts* 97–98 (2016) (unpublished manuscript) (on file with author) (posing a similar hypothetical in the context of someone randomly shooting in a theater and accidentally killing a hitman).

7. See Gideon Yaffe, *Criminal Attempts*, 124 *YALE L.J.* 92, 95 (2014).

persons, rather than also discussing actors who may be *reckless* or *negligent* regarding the infliction of such harm because— for purposes of unwitting justification—I believe the moral issues are the same.⁸ Similarly, and for the same reason, I confine myself to actors who do not *know* that inflicting harm is necessary to protect themselves or others from wrongful harms or lesser evils, rather than also discussing actors who are *not conscious of the likelihood* that inflicting such harm is necessary because, again, I believe the normative issues are the same.

Like Steinhoff, I think it is useful to have paradigmatic hypotheticals of unwitting justification in mind. So I am happy to accept his *Jill and Earl* hypothetical, where Jill shoots and kills Earl for entirely malicious reasons, unaware that doing so is necessary to prevent Earl from imminently and wrongfully killing an innocent third person:

A. *Jill and Earl*

Jill always wanted to murder Earl, whose nose she does not like. She sees him sitting in a restaurant, draws her gun, and shoots him dead. Unbeknownst to her, however, Earl, in turn, was just about to murder an innocent business rival who was also in the restaurant, and shooting Earl dead was the only way of preventing this.⁹

Jill and Earl is useful because it illustrates something that is true of all unwitting justification cases: regardless of whether unwittingly justified actors are guilty of completed crimes of harm, such as murder, they nonetheless possess the elements of impossibility attempts to commit such

8. Morally, a person is culpable for attempting to harm another if, while purposefully, knowingly, recklessly, or negligently disregarding the legitimate interests of the other, he engages in conduct that would violate the other's interests if the facts were as he mistakenly believes them to be. To be sure, as a matter of law, a majority of jurisdictions confine culpability for attempt to persons who act with purpose or knowledge. See MODEL PENAL CODE AND COMMENTARIES § 5.01 (AM. LAW INST., Official Draft and Revised Comments 1985). *But cf.* COLO. REV. STAT. § 18-2-101 (West 2018) (“A person commits criminal attempt if, acting with the kind of culpability otherwise required for commission of an offense, he engages in conduct constituting a substantial step toward the commission of the offense.”). However, the majority of jurisdictions take that position for administrative reasons—not because morals dictate it. And even then, many such jurisdictions that refuse to use laws against *attempt* to punish persons who engage in reckless but otherwise harmless conduct nonetheless punish such persons under laws against *reckless endangerment*. See MODEL PENAL CODE AND COMMENTARIES § 5.01 cmt. 2 (AM. LAW INST., Official Draft and Revised Comments 1985).

9. Steinhoff, *supra* note 6, at 138 (emphasis added).

crimes. This is true independently of whether jurisdictions define justification subjectively or objectively. To illustrate, let us assume that two identical cases arise—Steinhoff’s original *Jill and Earl*, and an identical case, *Jill-2 and Earl-2*. The difference between the two cases is that *Jill & Earl* arises within Jurisdiction X that defines justification subjectively—and thus treats Jill as culpable for murder—while *Jill-2 & Earl-2* arises within Jurisdiction Y that defines justification objectively—and thus does not treat Jill-2 as culpable for murder. Let us further assume that both jurisdictions define murder as “the unlawful killing of another human being with malice aforethought,”¹⁰ *unlawful* being a stand-in for *unjustified*, and that both jurisdictions adhere to the Model Penal Code (MPC) definition of impossibility attempts to commit crimes of harm, as follows:

B. Jurisdictions X and Y

Murder. A person is guilty of murder if the person unlawfully kills another human being with malice aforethought.

Attempt. A person is guilty of an [impossibility] attempt to commit a crime if, acting with the kind of culpability otherwise required for commission of the crime [and,] when causing a particular result is an element of the crime, [the person] does or omits to do anything with the purpose of causing or with the belief that it will cause such result without further conduct on his part¹¹

Definitions. *Unlawful* means without justification, as *justification* is otherwise defined in this jurisdiction.

It follows from the facts as stipulated that, regardless of whether actors like Jill and Jill-2 are in Jurisdiction X—which defines justification subjectively—or in Jurisdiction Y—which defines it objectively—they both satisfy the two elements of attempted murder under the MPC.¹² Thus, they both possess the culpability otherwise required for murder because, having acted with the *mens rea* of purpose regarding both the *harm* element of killing of a human being and the *circumstance* element of doing so unlawfully, they both acted with malice aforethought.¹³ Stated

10. 18 U.S.C. § 1111 (2012).

11. MODEL PENAL CODE AND COMMENTARY § 5.01(1) (AM. LAW INST., Official Draft and Revised Comments 1985) (addressing impossibility attempts to commit crimes of harm).

12. *See id.*

13. It might be argued that it was not Jill-2’s purpose to kill Earl-2 unjustifiably. However, *purpose* with respect to circumstance elements such as *lack of justification* means knowledge or belief; and Jill-2 did, indeed, believe that she was killing Earl without justification. *Cf.* MODEL PENAL CODE § 2.02(2)(a) (AM. LAW INST., Official Draft 1985) (“A person acts purposely with respect to a material element of an offense when . . . if the

differently, they both acted with malice because, if each had *succeeded* in unlawfully killing Earl—as, indeed, Jill succeeded, in that she killed Earl in Jurisdiction X that defines justification subjectively—neither of their convictions for murder would have failed for lack of *mens rea*. Both Jill and Jill-2 also satisfy the second element of attempt under the MPC because they both did something with the belief that it would cause Earl and Earl-2’s unlawful deaths “without further conduct on [their] part.”¹⁴

If the foregoing analysis is correct, it means that the dispute between objectivists and subjectivists is not about *whether* such actors should be punished but about *what*, as between attempt and completed crimes of harm, they are most appropriately punished *for*. To be sure, if impossibility attempts carried the same penalties as their respective crimes of harm, little would turn on the choice between subjective and objective definitions of justification. In practice, however, jurisdictions everywhere punish at least some impossibility attempts less severely than completed crimes of harm. The dispute between subjectivists and objectivists thus comes down to something that both sides strangely ignore; namely, the *reasons why* jurisdictions punish completed crimes more severely than impossibility attempts.¹⁵ The task, therefore, is, first, to identify those reasons and, then, to apply them to unwittingly justified actors to determine if the reasons militate in favor of punishing them for completed crimes of harm or for crimes of attempt.

III. WHY HARMFUL RESULTS MATTER

Criminal law theorists disagree about why resulting harms matter. Despite their disagreements, theorists tend to concur on two things: (1) no one should be punished in excess of what he deserves; and (2) the state cannot justly punish one deserving offender less severely than another, equally-deserving offender unless it does so for consequentialist reasons—deterrence, an offender’s dangerousness, or administrative economy. At the same time, theorists differ regarding whether criminal desert is solely a function of

element involves the attendant circumstances, he is aware of the existence of such circumstances or he believes or hopes that they exist.”). Indeed, if belief did not suffice with respect to attempts to commit crimes with circumstance elements, no one could ever be guilty of an impossibility attempt.

14. MODEL PENAL CODE AND COMMENTARIES § 5.01(1)(b) (AM. LAW INST., Official Draft and Revised Comments 1985).

15. See generally Michael Davis, *Why Attempts Deserve Less Punishment than Complete Crimes*, 5 L. & PHIL. 1 (1986).

the elements of attempt—possession of a guilty mind and willingness to act on it—or whether criminal desert is also a function of resulting harm. As a consequence, theorists who take desert to be solely a function of the elements of attempt—“anti-harm desert theorists”—believe it is unjust to grade punishment on the basis of resulting harm, unless doing so can be justified on consequentialist grounds—which they tend to deny.¹⁶ In contrast, theorists who take desert to be a function of resulting harm—“pro-harm desert theorists”—find differential punishment justified independently of consequentialist grounds.¹⁷

The aforementioned theories both have advantages and disadvantages. Pro-harm desert theories have the advantage of explaining the nearly universal criminal law practice of taking resulting harms into account. But they have the disadvantage of entailing the counterintuitive proposition that persons deserve to be punished for consequences that are not wholly within their control. In contrast, anti-harm desert theories have the advantage of basing desert on actions that are wholly within a person’s control—possession of a guilty mind and willingness to act on it. But they have the disadvantage of departing from the nearly universal practice of taking resulting harms into account in grading punishments.

Happily, an alternative theory exists that combines the advantages of pro-harm and anti-harm desert theories while avoiding their disadvantages. It is a theory that Plato advanced 2300 years ago in *The Laws*.¹⁸ Like anti-harm desert theorists, Plato takes the position desert is measured by attempts to commit crimes, regardless of outcome.¹⁹ According to Plato, actors who succeed in committing completed crimes and actors who fail while committing impossibility attempts are identical in the punishment they deserve.²⁰ They both deserve the same punishment because negative desert is complete when, being possessed of a guilty mind, an actor does everything he believes is necessary to bring the respective harm about, regardless of what thereafter

16. See generally, e.g., Andrew Ashworth, *Criminal Attempts and the Role of Resulting Harm Under the Code, and in the Common Law*, 19 RUTGERS L.J. 725 (1988); Sanford H. Kadish, *The Criminal Law and the Luck of the Draw*, 84 J. CRIM. L. & CRIMINOLOGY 679 (1994); Stephen J. Schulhofer, *Harm and Punishment: A Critique of Emphasis on the Results of Conduct in the Criminal Law*, 122 U. PA. L. REV. 1497 (1974).

17. See, e.g., R. A. DUFF, CRIMINAL ATTEMPTS 116–27, 327–47 (1996); R. A. DUFF, INTENTION, AGENCY AND CRIMINAL LIABILITY: PHILOSOPHY OF ACTION AND THE CRIMINAL LAW 184–92 (1990); R. A. Duff, *Acting, Trying, and Criminal Liability*, in ACTION AND VALUE IN CRIMINAL LAW 75, 75–106 (Stephen Shute, John Gardner & Jeremy Horder eds., 1993); Michael S. Moore, *The Independent Moral Significance of Wrongdoing*, 5 J. CONTEMP. LEGAL ISSUES 237 (1994).

18. See PLATO, THE LAWS IX: 876–77 (Trevor J. Saunders trans., Penguin Books 1970).

19. *Id.*

20. *Id.*

transpires.²¹ Thus, Plato writes, with respect to a malicious actor, A, who attempts to wrongly kill another person, B, A's criminal desert is complete at the point at which A *fully acts on his intent to kill* and is not a function of harms that may or may not thereafter occur.²² At the same time, Plato also takes the pro-harm desert position that, as between attempters who succeed in inflicting harm and attempters who fail, attempters who fail should be punished less severely than attempters who succeed, albeit not for consequentialist reasons.²³

How, then, does Plato manage to combine anti-harm views of desert with pro-harm views of differential punishments based on harm? He does so by rejecting one of the tenets on which pro-harm and anti-harm desert theories concur. While Plato agrees that persons should not be punished in excess of desert, he parts company with pro- and anti-harm desert theorists by rejecting the view that consequentialist grounds are the only reasons for punishing one equally-deserving offender less severely than another.²⁴ He argues instead that, although attempters who succeed in committing grievous crimes of harm should receive the full measure of the deserved punishment, attempters who fail should be punished less severely than they deserve for reasons independent of deterrence, dangerousness, or economy.²⁵

The key to Plato's argument is his rationale for reducing penalties for failed impossibility attempters. Society should reduce penalties for failed impossibility attempts, he says, not out of sympathy or pity for failed attempters or because society believes them to be less deserving of blame and suffering.²⁶ Instead, he notes, society should reduce penalties for failed attempts on *its own* account²⁷: society should reduce such penalties because society itself experiences—and *ought* to experience—different emotional reactions to failed attempts than to successful attempts, and because society's public punishments ought to *express* those different reactive emotions.²⁸ Successful attempts inflict wrongful harms on a community and, in doing so, they evoke—and rightly evoke—strong moral emotions, including emotions

21. *Id.*

22. *See id.*

23. *Id.*

24. *See* Peter Westen, *Why Criminal Harms Matter: Plato's Abiding Insight in the Laws*, 1 CRIM. L. & PHIL. 307, 316 (2007).

25. *Id.*

26. *Id.*

27. *Id.*

28. *Id.*

of regret that events transpired as they did and abiding grief over the wrongful harms that innocent persons suffered.²⁹

In contrast, failed attempts have the opposite effect. Failed attempts have the effect of sparing a community from wrongful harm and, in doing so, they evoke—and rightly evoke—emotions of profound relief that events did not transpire as offender envisaged them.³⁰ Punishment, in turn, is a vehicle by which a community expresses its reactive moral emotions regarding criminal conduct.³¹ A community that feels the same degree of righteous indignation when it is favored by fortune as when it is disfavored by fortune is a community with perverse moral emotions, at least when serious harms like homicide are at issue.³² And a community that expresses the same degree of resentment in the punishments it imposes is a community that risks offending the moral order.³³

IV. APPLYING PLATO'S RATIONALE FOR RESULTING HARM TO UNWITTING JUSTIFICATION

Plato confines his discussion of resulting harm to a malicious actor who tries but fails to kill an intended victim, and, hence, Plato does not address unwittingly justified actors like Jill in *Jill and Earl* who *do* kill their targets.³⁴ Nevertheless, I believe that it follows from Plato's reasoning that, for purposes of grading punishments on the basis of resulting harm, nearly all unwittingly justified actors, if not all of them, are analogous to malefactors who miss their targets—not to malefactors who hit their targets.

I shall divide unwitting justification cases into three sets: (A) cases in which society is positively relieved by the resulting harms that actors bring about; (B) cases in which society, although not positively relieved by resulting harms, feels no regret either; and (C) cases, if such cases exist, in which society does, indeed, regret resulting harms, despite their being justified.

A. Positive Relief

Some instances of unwitting justification cause societies to experience positive relief. To illustrate in the context of self-defense, consider *Jill and Earl*.³⁵ As previously discussed, Jill did everything she thought was necessary

29. *Id.*

30. *Id.*

31. *See id.*

32. *See id.*

33. *See id.* at 316–17.

34. PLATO, *supra* note 18. For *Jill and Earl*, see *supra* pp. 423–24.

35. *See supra* pp. 423–24.

to kill Earl and to kill him unlawfully.³⁶ If Plato is correct, Jill fully deserves just as much punishment as she would have deserved had she succeeded. However, jurisdictions are not morally obliged to give all classes of offenders the full measure of suffering they deserve for wrongdoing. Rather, jurisdictions are free to withhold a portion of deserved punishment when, for reasons that owe nothing to an actor's efforts, they are relieved by the results of an actor's unwitting conduct and, being relieved, are naturally—and even rightly—less disposed to exact the full pound of flesh that is their due. By that measure, Jill is identical to Plato's malicious actor. Plato's malicious actor and Jill both tried to kill a human being who, as far as they knew, was no threat to them or anyone else. Yet, despite their intentions, Plato's malicious actor and Jill both produced results that, far from leaving their respective communities in grief, left their communities feeling relieved. Plato's malicious actor's community is relieved because, in failing to kill his target, Plato's malicious actor unwittingly spared the life of an innocent person whom he would otherwise have wrongfully murdered. Jill's community, in turn, is relieved because, in hitting her target, she stopped Earl from killing an innocent person—Earl's business rival whom Earl would otherwise have wrongfully murdered. To be sure, Jill differs from Plato's malicious actor in that Jill committed a homicide, while the malicious actor did not. However, just as an otherwise omniscient community would wish *ex ante* that Plato's malicious actor not kill his target, an otherwise omniscient community would wish *ex ante* that Jill kill Earl because, given the circumstances, the alternative of Earl's killing an innocent person was morally worse.

The same can also be true in choice-of-evil contexts. Assume, for example, that Jack maliciously sets fire to Jane's cornfield, while being wholly unaware that a devastating forest fire is approaching. Jack's fire creates a much-needed fire break. Absent Jack's fire, the forest fire would destroy both Jane's cornfield and their nearby town. Assume, too, that if the town's fire fighters had deliberately set Jack's fire to create a fire break in the face of the oncoming blaze, they would have been feted as heroes for choosing the lesser evil. Like Plato's malicious actor, Jack is just as culpable as a successful arsonist because Jack did everything he thought was necessary to produce the evil of arson, all things considered. However, Jack's community may rightly do what it generally does with impossibility attempts: it may give Jack less punishment than he deserves because, far from being

36. See *supra* pp. 423–24.

aggrieved at what Jack did, the community is and ought to be relieved at what he did, all things considered.

B. *Lack of Regret*

Some instances of unwitting justification leave societies without feelings of positive relief, and, yet, without feelings of regret as well. To illustrate in the context of unwitting self-defense, consider the following variation on Jill and Earl:

1. *Jill-3 and Earl-3*

Jill-3 and Earl-3 are both NRA members who, for reasons of self-defense, are always armed. They are also bitter enemies. They parted ways years earlier and never expected to see one another again. Due to happenstance, however, they find themselves face to face with one another. Each decides at the very same instant to shoot the other, and each undertakes to do so. Jill-3, who succeeds in drawing her pistol slightly earlier than Earl-3, shoots and kills Earl-3 just as he is about to pull the trigger of a revolver aimed at Jill-3's heart.

Jill-3 and Earl-3 differs from Steinhoff's *Jill and Earl* in two respects.³⁷ First, while it is clear that Jill acted in objective self-defense by virtue of killing someone who was wrongfully threatening an innocent third party, it is less clear that Jill-3 was permitted to kill Earl-3. The difficulty arises from the simultaneity of their actions. Jill-3 cannot have been objectively permitted to kill Earl-3 by virtue of his posing a *wrongful threat* to her, because Jill-3 and Earl-3 were identically situated; and, if Earl-3 posed a wrongful threat to Jill-3, Jill-3 must also have posed a wrongful threat to Earl-3, thereby *depriving* Jill-3 of permission to kill him. By the same token, however, Jill-3 cannot have posed a wrongful threat to Earl-3 that permitted him to kill her because, again, the two actors were identically situated; and, if Jill-3 posed a wrongful threat to Earl-3, Earl-3 must have posed a wrongful threat to her, thereby depriving himself of permission to use lethal force against her. Nevertheless, Jill-3 and Earl-3 remain objectively permitted to use lethal force against one other to protect themselves, albeit not because either was a wrongful threat to the other. They are objectively permitted to do so because all actions are either permissible or wrongful; because the state should not punish a person for producing a wrongful harm unless he actually causes produces it; and because, given the paradoxical

37. See Steinhoff, *supra* note 6, at 138.

effects of simultaneity, the state cannot declare Jill-3 to have caused a wrongful harm without contradicting itself.

Second, Jill-3 differs from Jill because, while society is positively relieved that Jill killed Earl, society is not relieved that Jill-3 killed Earl-3 because the alternative would have been morally identical—Earl-3 killing Jill-3. Nevertheless, despite not being relieved by Jill-3's conduct, society does not regret it either. To be sure, society regrets that Jill-3 and Earl-3 ended up confronting each other, and society regrets that they both reacted in anger by trying to kill each other. But, given that those circumstances occurred, society does not regret the harm that resulted from Jill-3's conduct because society realizes that, if Jill-3 had not killed Earl-3, Earl-3 would have killed her.

Society's lack of regret at Earl-3's death is significant because it bears upon the distinction society draws in penalties between impossibility attempts and crimes of resulting harm. Like Plato's society, societies today generally reserve their highest penalties for persons who bring about the material harms that societies seek to prevent by means of punishment and regret when they occur. By the same token, societies today generally impose lesser penalties on persons who, in failing to bring such harms about, leave societies with no emotions of regret or grief under the circumstances. Jill-3 *tried* to bring about the harm of unlawful homicide, and she should be punished for attempting to do so. But she failed because, despite being unaware that she was doing so, she permissibly killed someone who otherwise would have killed her.

The same is true of some actors who choose the lesser of two evils. However, it is important to make two points beforehand. First, and contrary to what the MPC seems to imply,³⁸ if the lesser-of-two-evils principle is morally just, the metric by which it measures harms and evils cannot be a utilitarian one. That is, it is not a metric that asks, "As between harm X and harm Y, which harm causes less total unhappiness among victims, considered as a group?" A utilitarian metric of that kind would justify actions that are patently immoral—involuntarily harvesting a healthy person's body parts to save five dying patients. Rather, if the lesser-of-two-evils principle is truly just, the metric by which it measures harms and evils must be a moralized and contextualized one that takes multiple factors into considerations: whether an actor is inflicting morally wrongful harm as opposed to a harm that is not morally wrongful, whether an actor's conduct consists

38. See MODEL PENAL CODE § 3.04(1) (AM. LAW INST., Official Draft 1985).

of an action as opposed to an omission, and the manner in which harms are brought about, for example, switching a trolley to another track as opposed to pushing a fat man onto the tracks.³⁹

Second, if it is to be just, the lesser-of-two-evils principle cannot be confined to justifying the *lesser* evils but must also justify all that are *not greater evils*. To be sure, this caveat is immaterial if all evils are either lesser or greater than one another. However, if some evils are neither greater nor lesser but instead equal to one another in gravity, it would be fundamentally unjust to punish a person who has no choice but to opt for one or the other. It would be fundamentally unjust because, in addition to punishing a person who had no choice but to inflict an equal evil, it would mean punishing a person for bringing about a harm that, by definition, the state is indifferent about wishing to prevent under the circumstances.⁴⁰

With these two caveats in mind, consider a hypothetical Steinhoff adopts from Jonathan Quong:

2. *Man on the Track*

There is a runaway trolley whose brakes have failed headed down a track where your child is trapped and will be killed by the trolley. Fortunately there is a side track onto which the trolley can be diverted, but there is one man trapped on this side track, and he will be killed if you divert the trolley.⁴¹

Steinhoff believes, and Quong agrees, that the parent is permitted to divert the trolley in order to save his child, even though it means killing the stranger on the side track.⁴² Now consider the following variation:

3. *Unwitting Parent in “Man on the Track”*

There is a runaway trolley whose brakes have failed headed down a track where a mother’s child is trapped and will be killed by the trolley. The mother is unaware that the runaway trolley puts anyone’s life in danger,

39. For the proposition that the metric for measuring evils for purposes of necessity cannot be based upon net social utility, see JOSHUA DRESSLER, *CASES AND MATERIALS ON CRIMINAL LAW* 565 (5th ed. 2009); Alan Brudner, *A Theory of Necessity*, 7 *OXFORD J. LEGAL STUD.* 339, 341–44 (1987); Edward M. Morgan, *The Defence of Necessity: Justification or Excuse?*, 42 *U. TORONTO FAC. L. REV.* 165, 170 (1984); John T. Parry, *The Virtue of Necessity: Reshaping Culpability and the Rule of Law*, 36 *HOUS. L. REV.* 397, 414, 438 (1999).

40. Peter Westen & James Mangiafico, *The Criminal Defense of Duress: A Justification, Not an Excuse—and Why It Matters*, 6 *BUFF. CRIM. L. REV.* 833, 883–88 (2003).

41. Steinhoff, *supra* note 6, at 167 (adopting the *Man on the Track* hypothetical initially proposed in Jonathan Quong, *Killing in Self-Defense*, 119 *ETHICS* 507, 512 (2009)).

42. *Id.*

much less her child's. However, she sees that if she diverts the trolley on to a side track where an enemy of hers is trapped, the trolley will kill him. The mother diverts the trolley, thereby killing her enemy, only to discover afterwards to her relief that doing so was necessary to save her child's life.

I do not know whether Steinhoff and Quong are correct that society does not regret—or, at least, ought not regret—the death of the innocent person in *Man on the Tracks*, given the alternative harm that would have occurred, although I am inclined to think they are. However, if Steinhoff and Quong are correct that society does not regret the resulting harm in *Man on the Track*, society cannot regret the resulting harm in *Unwitting Parent in "Man on the Track"* either, because, from an objective standpoint, the actions, resulting harms, and alternatives are the same. In both cases, a parent saves the life of a child on the track of a runaway trolley by diverting a runaway trolley into the fatal path of an innocent person who otherwise would have lived.

To be sure, society is indignant—and righteously indignant—toward the mother in *Unwitting Parent in "Man on the Track"* because of the harm she tried to bring about, that is, the harm of killing an innocent person where the alternative was no harm at all. And society is right to be indignant and to want the mother to suffer because it is a crime for one to do all one thinks is necessary to kill a person whose death society regrets under the circumstances, even if one does not ultimately succeed.⁴³ However, the crime of trying and failing to cause a homicide that society regrets under the circumstances is *attempted* murder, not murder.⁴⁴ Murder is reserved for homicides that society wishes had not occurred under the circumstances.

C. Regret

I have thus far addressed instances of unwitting justification that leave society with no regrets, that is, with no wish that persons had acted otherwise given the alternatives. However, instances of unwitting justification may occur that leave society with just such regret. Consider, for example, a hypothetical of Jeff McMahan's that Steinhoff summarizes as follows:

43. See MODEL PENAL CODE § 5.01 (AM. LAW INST., Official Draft 1985).

44. Compare *id.* (delineating the circumstances of an attempted crime), with *id.* § 210.2 (delineating the circumstances of a completed murder crime).

1. *Tactical Bombers*

There is a bomber crew of 5 people who have, in the context of a humanitarian intervention, the mission to bomb and destroy a military target. If they are successful, 100 innocent civilians in the state in which the intervention takes place will be saved, and there are no alternative means of saving them. However, 5 innocent villagers will be killed in this attack as a side-effect.⁴⁵

McMahan believes, and Steinhoff assumes, that the bombers' planned attack on the military target is both "proportionate and justified."⁴⁶ Steinhoff also believes, while McMahan does not, that the innocent villagers would be justified in defending themselves by firing on the bomber pilots.⁴⁷ With that in mind, consider the following variation:

2. *Unwitting Villagers*

The bombers in *Tactical Bombers* make an initial pass over the military target, dropping a small bomb that kills nearby animals. The 5 villagers are furious and decide at that very moment to retaliate. The bombers return in order to commence what, without the villagers' realizing it, is a bombing run that will destroy the target and kill the villagers. The villagers fire on the bombers, unwittingly saving their own lives but at the cost of 5 crew members and 100 innocent civilians whom the bombers would otherwise have saved.

I do not know whether Steinhoff is correct in thinking that the villagers in *Tactical Bombers* are permitted to attack the bombers to defend themselves, though I am inclined to think he is. However, even if the villagers in *Tactical Bombers* are both permitted and nonculpable in defending themselves, the villagers in *Unwitting Villagers* can, at the very least, be justly punished for attempted homicide. They satisfy the elements of attempted homicide because, while acting in the belief that they were unlawful combatants, they did everything they believed necessary to kill a bombing crew that they believed was no threat to themselves.

The more interesting question, however, is whether the villagers' culpability in *Unwitting Villagers* is confined to that of *attempted* homicide or whether it extends to the graver crime of *criminal* homicide. If Plato and others are correct, the villagers in *Unwitting Villagers* deserve just as much punishment

45. Steinhoff, *supra* note 6, at 187 (summarizing a hypothetical included in Jeff McMahan, *Self-Defense Against Justified Threateners*, in *HOW WE FIGHT: ETHICS IN WAR* 104, 104–05 (Helen Frowe & Gerald Lang eds., 2014)).

46. *Id.*

47. *Id.* at 188.

as unlawful combatants who impermissibly kill a justified bombing crew.⁴⁸ However, it does not follow that society ought to give both sets of offenders the full punishment they theoretically deserve. Whether society ought to do so depends upon whether and, if so, why society is justified in punishing impossibility attempts less severely than completed crimes of harm. If Plato is correct, society is justified in giving even heinous offenders less punishment than they deserve when, for reasons that are of no credit to themselves, they fail to bring about harms that society regrets, all things considered. By that measure, it can be argued that the villagers in *Unwitting Villagers* should not be given less punishment than they deserve because they arguably *did* bring about a harm that society regrets, all thing considered—the death of 105 persons in lieu of 5.

Now it might be thought that, because the villagers in *Unwitting Villagers* inflict the same harm as villagers in *Tactical Bombers*, and because the latter villagers are permitted to inflict the harm, the harm itself cannot simultaneously be permissible and regrettable, all things considered. But that is not so. Whether an act is permissible depends upon whether it is a criminal wrong, that is, an act that society wishes to prevent by means of criminal law. Whether an act is regrettable depends upon whether it is an act that society wishes had come out differently. A society can simultaneously take the position that the villagers' act in *Tactical Bombers* is not a criminal wrong and, yet, wish the villagers had not committed it. To illustrate, consider the following:

3. *Breaching a Levy to Save a Town*

Rising flood waters threaten to flood a downstream town and kill hundreds of sleeping residents. A resident of the town, on orders from the governor, commences to breach an upstream levy in order to divert flood waters to an adjacent flood plain, though doing so will flood a nearby farm and killing a family of three that lives there. A member of the family notices the threat to himself and his family and, armed with a long-range rifle, shoots at the town resident, hoping to hit his target and prevent the resident from breaching the levy.

I assume that, if Judith Thomson's man-at-the-switch is justified in diverting a runaway trolley and thereby killing one on a side track to save

48. See PLATO, *supra* note 18.

five, the town resident is justified in diverting the rising river and thus drown three to save hundreds.⁴⁹ I also assume that, if Steinhoff is correct that the villagers in *Tactical Bombers* may kill the bomber crew to defend themselves, the farmer may kill the town resident to prevent the resident from killing himself and his family.⁵⁰ On those assumptions, it follows that, if the farmer succeeds in hitting his target, the farmer will not be punished for homicide. However, it does not also follow that the society as a whole is indifferent as to whether the farmer hits or misses his target. On the contrary, the society can simultaneously (1) maintain that, because of agent-relative considerations, the farmer is permitted to do his best to hit his target and, yet, (2) fervently hope that the farmer misses his target.

V. STEINHOFF'S ARGUMENTS AGAINST OBJECTIVE JUSTIFICATION

Steinhoff criticizes objective justification on four independent grounds: (A) it misconstrues the linguistic meaning of *justification*; (B) it requires that, if unwittingly justified actors are acquitted of completed crimes of harm, they must also be acquitted of criminal attempts and thereby “get entirely off the hook”; (C) it makes justified defense practically impossible; and (D) it misconstrues morality by failing to provide persons with “action-guidance.”⁵¹

A. *Objective Justification Misconstrues the Linguistic Meaning of Justification*

Steinhoff claims that objective justification is inconsistent with linguistic usage regarding self-defense:

If someone jumps out of the window above me with the intention to crush me and I, unaware of this fact, intentionally open my umbrella because it has started raining, and thereby avert his attack, then hardly anyone would say that I *defended* myself. . . . A roof tile knocking out my would-be murderer on a windy night also saves me, but no ordinary speaker would seriously say that the roof tile ‘defended’ me.⁵²

Steinhoff may be correct about ordinary language. But it begs the question to assume that criminal law adopts lay usage as its own. If objectivists are correct, *self-defense* is a term of art in law for actions by persons that are necessary and proportionate to protect them from wrongful harm. By

49. See generally Judith Jarvis Thomson, *The Trolley Problem*, 94 YALE L.J. 1395 (1985).

50. See Steinhoff, *supra* note 6, at 187.

51. *Id.* at 159.

52. *Id.* at 138–39.

that measure, Steinhoff's action of opening his umbrella is, indeed, self-defense, even though laypersons would use another term for it.

B. Objective Justification Requires That Unwittingly Justified Actors Get Off the Hook Entirely

Steinhoff argues that, if objectivists are correct that unwittingly justified actors are not guilty of completed crimes of harm, it follows that actors like "Jill should get entirely off the hook."⁵³ Actors like Jill must be let off the hook, he says, because it is "incoherent" to acquit an actor on grounds of justification and, yet, simultaneously convict him or her of attempted harm based upon the same underlying act.⁵⁴

In Part II we saw that, within jurisdictions that define objectively, it is perfectly possible to frame criminal statutes in a way that renders Jill guilty of attempted murder and, yet, not guilty of murder.⁵⁵ Indeed, all crimes of harm within such jurisdictions can be framed in that way. Jurisdictions can render actors guilty of attempted harm and not guilty of completed harm by (i) defining crimes of harm as the causing of *unlawful* harms, (ii) defining *unlawful* objectively, and (iii) adopting MPC standards regarding impossibility attempts.

Steinhoff recognizes this. He recognizes that, as a matter of law, statutes can be drafted to inculcate actors of attempt while acquitting them of completed crimes on grounds of justification.⁵⁶ His argument is different. He argues that, if a homicide like Jill's is morally and legally justified, it is incoherent to adjudge her guilty of attempted homicide based upon the same act because that would mean that the same act is "justified *and* unjustified":

While it is indeed completely consistent, given a suitable definition of murder, to say that Jill's successful killing of Earl was only *attempted* murder, it is *not* consistent to say that Jill's killing of Earl was *both* justified *and* unjustified. Thus if Jill's killing of Earl was *attempted murder*, and we deem attempted murder to be unjustified and impose criminal (and moral) liability for it on the act's perpetrator, then we cannot simultaneously say that the killing of Earl was justified self-

53. *Id.* at 139.

54. *Id.* at 145.

55. *See supra* pp. 423–24. Compare MODEL PENAL CODE § 210.2 (AM. LAW INST., Official Draft 1985) (delineating the circumstances of a completed murder crime), *with id.* § 5.01 (delineating the circumstances of an attempted crime).

56. *See Steinhoff, supra* note 6, at 139–40.

defense. One and the same act cannot be simultaneously unjustified attempted murder and justified self-defense.⁵⁷

With due respect, I believe Steinhoff's notion of *unjustified attempt* is a contradiction in terms. It is a *category mistake* to think that criminal attempts, once complete, are acts to which justification and lack of justification apply. Justification and the lack of it apply to offenses that are framed in terms of *prima facie* harms. Justification, and the lack of it, comes into play after an actor inflicts or risks such a *prima facie* harm, and does so to determine whether the harm he inflicted—or believed he was inflicting—was a harm the state does, indeed, wish to prevent, all things considered. In contrast, criminal attempt is not a crime of *prima facie* harm. It is a crime of intent of which the so-called *act* of attempt is merely constitutive evidence.

The difference is significant because, in contrast to offenses of *prima facie* harm, where matters of justification come into play *after* an offense is committed, matters of justification come into play *before* an attempt is committed. They come into play to determine whether a person is guilty of attempt at all—and, specifically, to determine whether what the person intends to do is, indeed, a harm or risk of harm that the state seeks to prevent, all things considered. If what a person intends to do is justified, a person is not guilty of attempt; if what a person intends to do is not justified, the person is guilty of attempt, provided the person engages in a requisite act and has the requisite mens rea regarding lack of justification. Thus, contrary to Steinhoff, there is no such thing as a *justified attempt*. A person who possesses the requisite intent and acts on it is guilty and ought to be punished, and to ask thereafter whether the attempt is justified or unjustified is meaningless.

Steinhoff relies to the contrary on what he believes to be implications of comments by Paul Robinson. Paul Robinson has argued that persons who commit impossibility attempts inflict *intangible harm* on their targets, which I take to be the dignitary harm that attempters inflict by doing their best to cause intended targets tangible harm.⁵⁸ Steinhoff relies on Robinson's statement to argue that, if attempts do indeed inflict harm, they should be treated like crimes of *prima facie* harm and, therefore, once committed, be assessed in light of any alternative harms that would have occurred in their absence, to determine if the attempts are justified.⁵⁹ I take no position here on whether attempts inflict dignitary harm on their targets. But, even if they do, it is a category mistake to classify such dignitary harms as being

57. *Id.* at 140.

58. *See* Robinson, *supra* note 2, at 291.

59. *See* Steinhoff, *supra* note 6, at 141–45.

commensurate with the *prima facie* harm that completed crimes inflict. Completed crimes are framed *in terms of prima facie* harms, and they are made criminal to prevent the harms from occurring—unless, of course, committing such crimes prevents greater harms from occurring.⁶⁰ In contrast, crimes of attempt are not framed in terms of dignitary harms, and they are not made criminal to prevent dignitary harms from occurring. The law criminalizes attempts to punish persons for what they intend to do as revealed through their acts, not because of any dignitary harms they may inflict in the process. And, therefore, the law does not weigh such dignitary harms against alternative harms that would otherwise occur to determine if attempts are justified or not.

Steinhoff, at the 2016 San Diego conference on his work, proposed a hypothetical designed to show that criminal attempts, after they are complete, can be justified by the favorable consequences they produce. In reality, Steinhoff's hypothetical shows the opposite: attempt is a crime of intent that is complete when an actor forms an intent to do what is unjustified and acts on the intent, and nothing of consequence that thereafter transpires can exculpate him of doing so. Steinhoff's hypothetical goes as follows:

Imagine the sadistic villain Scarlet credibly and truthfully threatens poor Kevin as follows: "If you do not [try] hard to steal the Matisse from the museum, I will explode an atomic bomb in New York. And if you actually succeed in stealing the Matisse, I will do so too. So your only hope is that the attempt gets thwarted. But you know that I can see if you do not try hard." Kevin's best bet to steal the Matisse is by sneaking into the museum's post office and changing the address on the envelope in which the Matisse has been put in order to send it to the art restorer. Kevin does that. The only way to leave the museum is via a Star Trek transporter, which will transport him into an isolation chamber where he has to remain for a week without any possibility of escape or communication. The envelope, however, would only need a day to arrive at Kevin's house. Kevin intends to keep the Matisse if he gets it and New York is destroyed—which he hopes won't happen. Thus he really does *intend* to steal it although he *hopes* not to succeed. Moreover, once Kevin has been transported to the isolation chamber his *attempt* is completed. It is out of his hands now.⁶¹

Steinhoff further states that Kevin takes sufficient steps in fulfillment of his intent to demonstrate that he is committed to carrying it out.⁶² But

60. See MODEL PENAL CODE AND COMMENTARIES § 3.02 (AM. LAW INST., Official Draft 1985).

61. Uwe Steinhoff, *Replies* 55 SAN DIEGO L. REV. 469, 503 (2018).

62. *Id.* at 503–04.

before he succeeds in mailing the Matisse to his house, a third party intervenes to thwart the theft and spare New York City as a consequence.⁶³

Based upon these facts, Steinhoff reasons that (1) Kevin satisfied the elements of attempted theft because what he intended to—steal the Matisse, with its consequence of destroying New York City—was unjustified, and, yet, (2) Kevin's *criminal attempt* to steal it was ultimately justified and, hence, a basis for exculpation, given that it resulted in a thwarting of the theft and a negating of the bomb threat.⁶⁴ I think Steinhoff is mistaken on both counts.

Let us start with Steinhoff's conclusion that Kevin satisfied the elements of attempted theft because what Kevin intended to do—steal the Matisse—was unjustified. Whether a person is justified in committing a crime depends upon the relative gravity of the alternative evils that would otherwise occur.⁶⁵ If the crime the person inflicts is greater in evil than the alternative evil that would otherwise occur, he is unjustified in committing the crime. Otherwise, he is justified.⁶⁶ With respect to attempt, a person is guilty only if what he intends to do—including any crime he intends to commit—is unjustified by the alternatives that would otherwise occur. And with respect to theft, a person is guilty of attempted unjustified theft only if what he intends to do—including any theft he intends to commit—is unjustified by the alternatives that would otherwise occur.

Now consider the dilemma Kevin faced. Two slim possibilities existed for Kevin to avert an atomic blast: (1) to give third parties some time to thwart the theft by stalling, dissembling, and otherwise pretending to steal the Matisse without actually intending to do so, knowing that any effort to dissemble might cause Scarlet to immediately detonate the bomb; or (2) to give third parties some time to thwart the theft without risking an explosion in the meantime by meticulously taking the multiple necessary steps on his part to complete the theft, knowing that, if the steps were they not interrupted as he hoped they would be, Scarlet would detonate the bomb as a consequence. Given Kevin's grim alternatives, it can hardly be said that the latter alternative, which is what Kevin chose, was the wrong course of action under the circumstances. Because it was not the wrong course for a person in Kevin's position to take, Kevin, by virtue of intending

63. *Id.*

64. *Id.*

65. See MODEL PENAL CODE AND COMMENTARIES § 3.02 (AM. LAW INST., Official Draft 1985).

66. *See id.*

it, did *not* intend what was unjustified and, hence, was *not* guilty of attempted unjustified theft.⁶⁷

Steinhoff's second conclusion is based on the assumption that Kevin satisfied the elements of attempting to commit a theft that he knew to be unjustified. So let us assume, contrary to what I have argued above, that Kevin satisfied the elements of attempted theft. On that assumption, Steinhoff argues that because the consequences of Kevin's attempted theft were that the theft was thwarted and New York City was saved, the latter favorable consequences *justified* Kevin's criminal attempt and should exculpate him of what would otherwise have been an unlawful, criminal attempt.⁶⁸

Again, Steinhoff mistakenly conflates justification with a category of offense—attempt—to which it does not apply. Crimes of *prima facie* harm can be justified by their consequences because such crimes are framed in terms of harms that society seeks to prevent, and they therefore serve no purpose when an actor, by committing them, spares society even greater harm, all things considered. Crimes of attempt are not framed in terms of harm. They are crimes of intent—and, specifically, intent to inflict unjustified harm—and they are complete when such intents are formed and acted upon, and nothing that thereafter happens as a consequence can justify it.

To illustrate, assume that a shopkeeper shoots to kill a fleeing shoplifter, knowing that the law bars the use of deadly force to prevent a shoplifter from escaping. The shopkeeper misses, and the shoplifter remains unharmed. However, as a result of the shooting, the police arrive, notice a fire in a nearby apartment house, call the fire department, and save many families that would otherwise have burned to death. The shopkeeper's attempted murder clearly had favorable consequences that society welcomes given

67. I put aside an additional reason why Kevin is not guilty of attempted unjustified theft. Attempt is a crime of intent, and therefore, an actor is not culpable unless he possessed *mens rea* regarding lack of justification. Compare MODEL PENAL CODE AND COMMENTARIES § 3.04(2)(b) (stating that a person who tries and fails to kill another is not guilty of attempted murder if he believes that doing so is necessary to protect himself from death or grievous bodily harm, even if his belief is unreasonable), with *People v. Goetz*, 68 N.Y.2d 96, 110–16 (1986) (reasoning that a person who tries and fails to kill another in the actual belief that doing so is necessary to protect himself from death or grievous bodily harm is nevertheless guilty of attempted murder if his belief was unreasonable). Given that Kevin chose to steal—in lieu of pretending to steal—because he thought stealing maximized his chances that the theft would be thwarted, he clearly believed his action was justified under the circumstances. In addition, given that the course of action he chose *did* thwart the theft and spare New York City, he can hardly have been negligent in believing that stealing was preferable to pretending to steal.

68. Steinhoff, *supra* note 61, at 503–05.

the alternatives. But none of that exculpates the shopkeeper of attempted unjustified murder because his was a crime of intended unjustified action that was completed when he acted on it.

*C. Objective Justification Makes Justified Self-Defense
Practically Impossible*

Steinhoff believes that the “gravest [problem with objectivism] is that it [renders justification] practically impossible.”⁶⁹ Objectivism does so, he says, because it implies a *necessity* requirement that self-defenders can at best satisfy by “pure blind luck.”⁷⁰

To illustrate, Steinhoff revisits Jill’s killing of Earl. Steinhoff previously stipulated that it was “necessary” for Jill to kill Earl to prevent Earl from wrongfully killing a rival.⁷¹ Steinhoff later questions the stipulation, arguing that for objectivists, it is rarely, if ever, necessary for an actor like Jill to kill another in self-defense because, regardless of how unlikely and fortuitous such possibilities may be, objective possibilities nearly always exist for bringing about milder outcomes that would effectively prevent wrongdoers like Earl from causing wrongful harm.⁷² For example, Jill’s pulling the trigger the exact moment and in the exact direction necessary for her bullet to collide in midair with Earl’s bullet thereby renders Earl’s bullet harmless.⁷³ To be sure, actors like Jill quite likely lack the knowledge and skill needed to ensure such milder outcomes or even make them likely to occur. Yet such outcomes are nevertheless possible. And because they are nearly always possible, it is rarely if ever necessary for an actor like Jill to kill or seriously harm another to defend herself or others:

[I]t is a simple fact of physics that . . . if Jill had fired a shot at a certain time *t* while holding her weapon in a certain position *p* the bullet of her gun would have struck the bullet coming out of Earl’s gun in midair and deviated it in such a way as to save the life of Earl’s innocent rival. Of course, there is no way that Jill can know which time and position that is, but since we are talking about objective justification, this is irrelevant Jill could, objectively speaking, do all the things I have just described . . . and thus it was objectively unnecessary and unjustified for her to kill Earl. This holds for virtually all cases of self-defense with a firearm. (Similar things, of course, can be said for similar reasons about all other kinds of self-defense: There will have hardly ever existed a case of self-defense where, objectively speaking, the same defensive result could not have been achieved with a *milder* means.)⁷⁴

69. Steinhoff, *supra* note 6, at 159.

70. *Id.*

71. *Id.* at 138.

72. *Id.* at 147.

73. *Id.*

74. *Id.* at 148.

The theoretical existence of such milder possibilities has two consequences for objectivists, Steinhoff argues.⁷⁵ For one thing, it means that persons like Jill are rarely, if ever, objectively justified in defending themselves or others. For another, it means that self-defenders like Jill are actually culpable for the unjustified harms they inflict because, given that it is evident that such milder possibilities exist, self-defenders must be aware of them, and, if they are not, they are negligent for being unaware of them.⁷⁶

Steinhoff's argument suffers from two problems that, I believe, are fatal to his position: (1) his critique, if valid, applies as much to his own theory of justification as to objective justification, thereby producing the *reductio ad absurdum* that, regardless of which theory one adopts, justified self-defense hardly ever obtains; and (2) his critique is premised on an erroneous notion of *necessity* for purposes of objective justification.

*1. Steinhoff's Critique Applies as Much to His Own Theory of
Justification as to Objectivism*

Steinhoff's view of justification is *mixed* in that it is neither wholly objectivist nor wholly subjectivist in nature. Steinhoff takes the position that a self-defender is justified if, and only if, the self-defender "*knows that the objective conditions of justified self-defense are fulfilled [and that a self-defender has] 'knowledge' [if, and only if, he has an] epistemically justified . . . true belief.*"⁷⁷ An actor like Jill is not justified, therefore, unless she *believes* that killing Earl is the mildest thing she can do to prevent Earl from wrongfully killing his rival and it is also *true* that killing Earl is the mildest thing she can do to prevent Earl from killing his rival.

With Steinhoff's mixed theory in mind, let us return to Steinhoff's critique of objectivism. Steinhoff criticizes objectivism on the ground that it leads to the *reductio* that (i) no one is hardly ever justified in killing in self-defense, and (ii) nearly everyone who kills in self-defense believing it is necessary is culpable for thinking so.⁷⁸ Notice, however, that because Steinhoff takes the position that, to be justified, an actor must believe in necessity and his belief must be "true," Steinhoff's theory suffers from the same *reduction* as he attributes to objectivism.⁷⁹ Consider Jill. Jill believed

75. *Id.*

76. *Id.* at 149.

77. *Id.* at 157–58 (footnote omitted).

78. *Id.* at 149.

79. *Id.* at 158 (emphasis omitted).

it was necessary to kill Earl to prevent him from wrongfully harming his rival. However, Steinhoff has shown—and Steinhoff is correct—that measured by everything that it was theoretically possible for Jill to do, killing Early was *not* the mildest thing that Jill could have done to prevent Earl from wrongfully harming his rival, and Jill’s belief to the contrary was *not* true.⁸⁰ On Steinhoff’s own theory, therefore, it follows that Jill was not only *not* justified in killing Earl but also *culpable* in thinking otherwise.

Steinhoff has a response to this. He acknowledges that under his mixed theory, a self-defender is not justified unless, in addition to the defender knowing that the *objective* conditions of self-defense obtain, it is also *true* that those objective conditions obtain.⁸¹ Nonetheless, he argues that because *it is wrong* to define objective conditions to include any and all milder means by which defenders could in theory protect themselves, his theory—in contrast to objectivism—does not define objective conditions in that way.⁸² As I argue below, I agree with him that it is wrong to define the conditions of self-defense—and, particularly, to define necessity—to include any and all milder means by which defenders could in theory protect themselves.⁸³ Indeed, it is precisely because it is wrong that objectivism is not obliged to adopt it, any more than Steinhoff adopts it for purposes of his mixed theory.

2. Steinhoff’s Critique is Based on a Mistaken Notion of Necessity

Fortunately, there is a way escape Steinhoff’s *reductio*. His *reductio* is based upon a mistaken notion of necessity. Although Steinhoff may be right that it is theoretically possible for actors like Jill to do milder things to protect themselves than resort to homicide, he is mistaken that theoretical possibilities of that kind constitute the metric by which *necessity* is measured for purposes of objective justification.⁸⁴ The appropriate metric for objective justification is not an *ex post* measure of the mildest act that it is *theoretically possible* for Jill to have performed, however infinitesimally small such possibilities might be. It is an *ex ante* measure of probabilities as they existed at the time Jill acted. It is an *ex ante* measure of the mildest act that it is *likely* that Jill, given her individual physical and psychological capacities, *could have succeeded in bringing about* to prevent the harm. Anything less would oblige self-defenders,

80. *Id.* at 147–48.

81. *Id.* at 157–58.

82. *See id.* at 150; Steinhoff, *supra* note 61, at 510.

83. Steinhoff, *supra* note 6, at 147–48.

84. *See id.*

under penalty of punishment, to either sacrifice themselves to wrongful aggression or perform actions they have no likelihood of accomplishing.

This does not mean that the metric is “epistemic”⁸⁵ in being subjective to the actor at issue. The metric is an objective standard in being based upon society’s considered judgment regarding the *ex ante* probabilities that existed at the time the actor acted, regardless of whether the actor himself was, or was not, aware of the probabilities at the time.⁸⁶ Nor does the metric reduce the metric to a “reasonability requirement”⁸⁷ in being based upon facts known to the actor at the time he acted. It is based upon such facts regarding *ex ante* probabilities as society is aware of at the time of judgment, regardless of whether they were or were not known by the actor at the time he acted. It means that, regardless of whether one is an objectivist or mixed-view adherent like Steinhoff, an actor need not elect a milder over a harsher means for defending himself, unless, in society’s judgment and based upon facts known to society at the time of judgment, an *ex ante* probability existed that the actor was likely to succeed in bringing about the milder means.

D. Objective Justification Fails to Provide Action Guidance

Finally, Steinhoff argues that objectivism fails the essential test of morality because it assesses right conduct by the results of an actor’s actions rather than by beliefs:

[A] purely objectivist account cannot provide proper action-guidance. To wit, if a proponent of this view is asked by someone: “I would like to go into a full theater and shoot randomly at the guests, is that permissible?”, then on pain of inconsistency the proponent would have to answer: “Well, that depends on who you will hit. If you hit people like Earl (or Hitler), it’s fine, otherwise not.” These answers, however, would certainly not be given by ordinary speakers, that is, by speakers who understand what terms like “permissible” and “justified” mean [W]hat we expect from morality is *action-guidance* [M]orality should be able to advise human beings as to how they should act under conditions of imperfect information

. . . .

. . . It is not permissible to go to a vegetable market in order to “randomly spray a large crowd of people with bullets,” not even if, accidentally, one shoots the

85. *Id.* at 147.

86. Kenneth W. Simons, *Self-Defense: Reasonable Beliefs or Reasonable Self-Control?*, 2 NEW CRIM. L. REV. 51, 57–58 (2008).

87. Steinhoff, *supra* note 6, at 150 (emphasis omitted).

“right” person. By acting in the irresponsible way the gunman acted, he might not have *violated* the *rights* of the innocent people in the market, yet he nevertheless showed insufficient respect for them . . . [A]n agent who engages in some act that is normally prohibited . . . must be required to have a *reasonable belief* that the objective justifying circumstances are fulfilled . . .⁸⁸

Steinhoff is right that morality should provide persons with guidance. But he is wrong that objective justification fails to do so, at least regarding Plato’s objectivist account. A Platonic objectivist envisages the very same action guidance as Steinhoff because, like Steinhoff, he assesses a person’s criminal desert by the wrong that the person *believes* or *ought to know* he is in fact committing, not by harms that result from his fully acting on his beliefs. A Platonic objectivist would give the vegetable-market shooter the very same warning that Steinhoff would give him. They would both warn him, “Do not try to kill people whom you believe or ought to know you are unjustified in killing.” The only difference between the two lies in what they would punish the shooter *for*. Steinhoff would punish the shooter for murder; in contrast, a Platonic objectivist would punish the shooter for attempted unjustified murder—that is, for trying but failing to kill persons whose deaths society regrets, all things considered.

VI. CONCLUSION

Uwe Steinhoff, being a mixed subjectivist and objectivist, maintains that a person has a defense of justification if, and only if, the person’s conduct is objectively justified, and the person believes his conduct is justified—thus denying a defense to a person who is unaware of it. I argue that Steinhoff’s criticisms of objective justification fail, in part, because he neglects to address the most persuasive account we possess regarding why and when resulting harms matter in criminal law.

Ultimately, the issue of unwitting justification has less practical significance than theoretical significance for what it reveals about criminal responsibility. It is often said that criminal law is wholly or largely designed to protect citizens from unjustified harms.⁸⁹ Unfortunately, the fact that criminal law largely or wholly *aims* to prevent unjustified harms can mislead one regarding the *content* of criminal offenses.

First, the emphasis on preventing unjustified harms can mislead one regarding the nature of paradigmatic criminal offenses. It can mislead one into thinking that the paradigm of which other criminal offenses are normatively derivative is a completed crime of unjustified harm such as

88. *Id.* at 154–57.

89. See H. L. A. HART, PUNISHMENT AND RESPONSIBILITY 8–11 (1968).

murder or rape. The reality, as Plato observed, is just the opposite.⁹⁰ Because criminal law prevents harm by means of punishment, and because punishment is just only if it is deserved, the paradigmatic criminal offense is one that is *constitutive of the elements of criminal desert*, namely, the crime of impossibility attempt. Impossibility attempts establish the baseline by which criminal desert is measured, whether actors succeed or fail in their wrongful undertakings.⁹¹ To be sure, attempts are *conceptually* derivative of offenses of unjustified harm and risk because attempts must necessarily make reference to such offenses. *Normatively*, however, impossibility attempts are anterior to crimes of unjustified harm and risk because impossibility attempts consist of the elements that establish an actor's culpability for such offenses. When society punishes completed crimes of grievous harm more severely than impossibility attempts, it is not because the former are more deserving of punishment than the latter. It is because society is materially aggrieved by the former and, being aggrieved, is indisposed to give them anything less punishment than the punishment they deserve.

Second, the emphasis on preventing unjustified harms can mislead one into thinking that all offenses consist in part of an *actus reus* of harm or risk—that is, an objective harm or risk that, unless justified by a countervailing harm or risk, is a wrong that must be prevented. However, the latter conception fails to account for the crime of attempt. Attempt does not consist of an *actus reus* as such, and, hence, once a criminal attempt is complete, it is incapable of justification. Attempt is essentially an offense of intent because the law's purpose in requiring an *act* of attempt is to ensure that an actor possesses an intent of a certain kind, namely, an intent that he is actually committed to carrying out.

Finally, the emphasis on preventing unjustified harms can mislead one regarding the nature of *mens rea*. It is commonplace to think that *mens rea* consists of culpability *regarding* the *actus reus* of offenses, much as the *mens rea* of murder consists of *malice aforethought* regarding acts of homicide. Again, however, this conception fails to account for the crime of attempt. Attempt does contain *mens rea* elements, to be sure. However, the *mens rea* of attempt is not directed to the *actus reus* of attempt. Quite the contrary, the *mens rea* of attempt *is* the crime of attempt, the so-called

90. See *supra* Section III.

91. A person who succeeds in committing a completed crime of harm or risk nonetheless possesses the elements of an impossibility attempt.

act of attempt being mere evidence that the actor is committed to effectuating his intent.