

Summer 2004

# Unknowingly Justified Actors and the Attempt/ Success Distinction

Larry Alexander

Follow this and additional works at: <http://digitalcommons.law.utulsa.edu/tlr>



Part of the [Law Commons](#)

---

### Recommended Citation

Larry Alexander, *Unknowingly Justified Actors and the Attempt/Success Distinction*, 39 *Tulsa L. Rev.* 851 (2013).

Available at: <http://digitalcommons.law.utulsa.edu/tlr/vol39/iss4/7>

This Article is brought to you for free and open access by TU Law Digital Commons. It has been accepted for inclusion in *Tulsa Law Review* by an authorized editor of TU Law Digital Commons. For more information, please contact [daniel-bell@utulsa.edu](mailto:daniel-bell@utulsa.edu).

## UNKNOWINGLY JUSTIFIED ACTORS AND THE ATTEMPT/SUCCESS DISTINCTION

Larry Alexander\*

In 1975, Paul Robinson, one of George Fletcher's star pupils and a recent graduate of the UCLA School of Law, wrote an article in the school's law review taking the position that "an act found to be beneficial, or at least not harmful, should be of no concern to the criminal law."<sup>1</sup> Robinson cites in support a statement of Macaulay:

When an act is of such a description that it would be better that it should not be done, it is quite proper to look at the motives and intentions of the doer, for the purpose of deciding whether he shall be punished or not. But when an act which is really useful to society, an act of a sort which it is desirable to encourage, has been done, it is absurd to inquire into the motives of the doer, for the purpose of punishing him if it shall appear that his motives were bad.

If A. kills Z. it is proper to inquire whether the killing was malicious; for killing is *prima facie* a bad act. But if A. saves Z.'s life, no tribunal inquires whether A. did so from good feeling, or from malice to some person who was bound to pay Z. an annuity; for it is better that human life should be saved from malice than not at all.<sup>2</sup>

Robinson spends a few pages discussing whether risky behavior is itself harmful despite its failure to produce the harm risked, and whether inchoate crimes and impossible attempts are harmful.<sup>3</sup> But the debates over whether risky conduct, inchoate crimes, and impossible attempts are harmful are not the debates Robinson wishes to engage. Robinson's interest is in whether an otherwise criminal act that is taken in circumstances that would amount to a legal justification for that act can be viewed as harmful in the sense that would justify criminal punishment. And Robinson concludes that it cannot.

What practical consequences follow from this seemingly obvious point that conduct that would otherwise be criminal but that occurs in circumstances that would legally justify it is not harmful in a sense meriting punishment? Robinson argues that a great deal follows. First, an act that is justified should not be

---

\* Warren Distinguished Professor, University of San Diego School of Law.

1. Paul H. Robinson, *A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability*, 23 UCLA L. Rev. 266, 267-68 (1975).

2. *The Works of Lord Macaulay*, vol. 7, 552-53 (Lady Trevelyan ed., Longmans, Green & Co. 1866).

3. Robinson, *supra* n. 1, at 268-71.

punished even if the actor negligently, recklessly, or deliberately created the conditions requiring the otherwise-criminal-but-now-justified act. Someone who starts a forest fire for the purpose of then setting a now-justified backfire on the victim's land should not be punished for the backfire. Rather, he should be punished for setting the original fire, the harm of which encompasses the now-justified backfire.<sup>4</sup> And because some criminal codes deny the justification defense to defendants who have created the necessity of the justified act, those codes are misguided and should be changed. They deter acts that are socially desirable and thus should be encouraged, not deterred. We want the defendant to set the backfire and avert even greater damage, even if he wrongfully created the circumstances that make the backfire socially desirable. He should be punished for his wrongful creation of the original fire, but not for setting the backfire.<sup>5</sup>

Robinson also criticizes those criminal codes that deny a justification defense to defendants who act with knowledge of the justifying facts but who are not motivated by them. Under some criminal codes, a defendant who sees that a backfire is justified but who is motivated to start one, not by the justifying concerns (averting as much damage as possible), but by the damage the backfire will cause to his enemy's property, cannot avail himself of a justification defense.<sup>6</sup> But surely that is a perverse result, as Robinson correctly argues. We want people to take the justified action regardless of their motives for doing so. Because they generally have no legal *duty* to take the justified act, we need to take advantage of all the extra-legal incentives we can get, including malicious and anti-social ones, to spur actors to take justified actions. And we surely do not want through the prospect of punishment to *deter* actors from taking justified actions whenever they lack any inclination to do so other than as a result of malicious and anti-social

4. *Id.* at 281-83.

5. Robinson erroneously accuses the Model Penal Code of making the mistake that other codes make because it allows punishment of the defendant who starts the justified backfire at the level of culpability that matches his culpability in starting the original fire. *Id.* at 282; see Model Penal Code § 3.02(2) (ALI 1985). Thus, under the Model Penal Code, the defendant who recklessly starts the original fire can be punished for "reckless arson" for setting the justified backfire. Robinson criticizes this result because it would deter the defendant from taking the socially desirable course of action and setting the backfire.

My own view is that Robinson has misread the Model Penal Code on this point. In part, it is because the Model Penal Code's wording makes it appear that the defendant who recklessly starts the original fire and then starts the now-justified backfire is punishable for the *latter* at the level of recklessness. I believe, however, that the Code's intent is to punish the defendant for reckless arson, and it does not distinguish—as it should—the original reckless arson from the subsequent justified arson. What is clear is that the Code does not wish to deter the justified arson. And indeed, were the defendant to sit on his hands after recklessly starting the fire, he would be guilty of a *knowing* failure to avert the harms of arson *after having acquired, through his recklessness in starting the fire, an affirmative duty to act*. If this would constitute a more serious crime than reckless arson, punishing the defendant at the level of recklessness for the original fire will not deter him from starting the backfire.

(In cases where the defendant *deliberately* starts the original fire, but now has the opportunity to minimize its damage by starting a backfire, the incentive to do the latter must come from either treating the deliberate failure to set the backfire as a second crime, one separate from the original arson, or from making the punishment for arson turn on the magnitude of the damage risked. For further discussion of related issues, see Larry Alexander & Kimberly D. Kessler, *Mens Rea and Inchoate Crimes*, 87 J. Crim. L. & Criminology 1138, 1183-93 (1997).

6. See Robinson, *supra* n. 1, at 284-87.

motives. The executioner does not commit murder merely because he volunteers for the job out of personal animosity for the condemned rather than a sense of civic duty.<sup>7</sup> And the same should be true of the defendant who starts an otherwise justified backfire solely out of hatred for the landowner.

Robinson's last point in the article is that if an act would otherwise be justified, it should not lose its justified status merely because the defendant was unaware of the justifying circumstances at the time he acted.<sup>8</sup> The point here is different from the preceding one about motive. For one can be aware of the justifying facts and yet be motivated, not by them, but by other concerns. Rather, the point under consideration here is whether the defendant who commits what is otherwise a crime in circumstances that would legally justify his conduct, *but who is not only unmotivated by the justifying facts, but is unaware of them*, may avail himself of the justification defense.

On this point, the Model Penal Code, almost all the extant case law from Anglo-American jurisdictions, and the treatise writers are in agreement: One who commits an otherwise criminal act ignorant of justifying facts may not later invoke those facts as a defense to the crime.<sup>9</sup> Robinson demurs, however. He believes justification defenses should be available to the unknowingly justified actor.<sup>10</sup> The actor has, of course, *attempted* to commit a crime and perhaps should be punishable as an attempter.<sup>11</sup> But like an attempter, the unknowingly justified actor has committed no social harm and should not be punishable for more than an attempt.<sup>12</sup>

George Fletcher responded to Robinson's article, with emphasis on the unknowingly justified actor.<sup>13</sup> Fletcher's response is characteristically subtle and

7. See *id.* at 286.

8. See *id.* at 288-91.

9. See *id.* at 288-89. A representative statement is that of LaFave and Scott, which Robinson quotes:

To have the defense of necessity, the defendant must have acted with the intention of avoiding the greater harm. Actual necessity, without the intention, is not enough. If *A* kills his enemy *B* for revenge, and he later learns to his happy surprise that by killing *B* he saved the lives of *C* and *D*, *A* has no defense to murder. In other words, he must believe that his act is necessary to avoid the greater harm.

Wayne R. LaFave & Austin W. Scott, Jr., *Handbook on Criminal Law* 386 (West 1972) (footnote omitted). The Model Penal Code begins each of its justification defenses with a reference to a defendant's *belief* in the existence of justifying facts. See e.g. Model Penal Code §§ 3.02(1), 3.03(3), 3.04(1)-(2), 3.05(1), 3.06(1), 3.06(3), 3.07(1)-(5), 3.08(2), 3.08(5)-(7).

10. Robinson, *supra* n. 1, at 289-91.

11. As the actor would under Model Penal Code § 5.01(1)(a), for example, which treats as an attempted crime conduct purposely undertaken "with the kind of culpability otherwise required for commission of the crime" and which "would constitute the crime if the attendant circumstances were as [the defendant] believes them to be . . ."

12. Robinson—mistakenly, in my opinion—suggests that the actor who knows the justifying facts but is motivated not by them, but by malice, should also be guilty of an attempt. Robinson, *supra* n. 1, at 290-91. But this would make the malicious executioner an attempter, which would be an undesirable outcome. And more generally it would, through the prospect of attempt liability, deter all justified conduct not motivated by the justifying facts or which the defendant fears might be found to be not motivated by those facts.

13. See George P. Fletcher, *The Right Deed for the Wrong Reason: A Reply to Mr. Robinson*, 23 UCLA L. Rev. 293 (1975).

erudite. He applauds Robinson's "legalist," non-instrumentalist approach to criminal law theory, and he concurs with Robinson that if the existence of some fact is required by the definition of the crime, its nonexistence should preclude conviction regardless of the defendant's awareness of its nonexistence and whether or not the defendant is a bad character in need of reform and incapacitation.<sup>14</sup>

Despite his approval of Robinson's legalistic approach to the unknowingly justified actor, Fletcher refuses to endorse Robinson's bottom line. He points out that Robinson's approach conflates the facts that go to the definition of the crime with those that establish the absence of a justification.<sup>15</sup> For Robinson, every crime, when fully defined, has the form of "Do not do X in circumstances not-J," where "not-J" means "not justifying X."<sup>16</sup> When the unknowingly justified actor does X, he does so in circumstances "J" rather than "not-J." Therefore, he has not committed the crime.

Fletcher, however, thinks that there is a distinction between definitions of crimes and justifications, and that the latter do not always function merely as negative elements of the former. Definitions and justifications are relevantly different. Definitions, for example, cannot be overly vague, whereas justifications can.<sup>17</sup> And legislative control over definitions is much more important than legislative control over justifications.<sup>18</sup>

Having pried apart definitions and justifications, which Robinson had arguably conflated, Fletcher then concludes his response by rather briefly turning to the problem of the unknowingly justified actor. His conclusions are tentative. Some unknowingly justified actors—for example, those unaware that the victim has consented to their conduct—should be regarded as having done no wrong.<sup>19</sup> Others, however, such as one who attacks a victim who he is unaware is about to attack him or others, have committed wrongs.<sup>20</sup> The attack on such a victim might be "just," but it is not "justified."<sup>21</sup>

14. *Id.* at 301-05.

15. *Id.* at 309-10.

16. Interestingly, the Model Penal Code treats the absence of justification as an element of every crime, and one to be proved by the prosecution beyond a reasonable doubt. Model Penal Code § 1.13(9). This would seem to be at least in tension with, if not inconsistent with, its requirement that one must know of the justification at the time one acts to be able to invoke it in defense. It is not inconsistent with the requirement because the Model Penal Code's scheme could be rendered as "Do not do X in circumstances not-J, where not-J includes belief that not-J." Nonetheless, the Model Penal Code's treatment of justifications as negative elements of crimes seems more in line with Robinson's approach to the unknowingly justified actor than with its own.

17. See Fletcher, *supra* n. 13, at 313-15.

18. See *id.* at 316-18.

19. *Id.* at 319.

20. See *id.* at 320.

21. *Id.* at 320-21.

In 1978, in his magisterial *Rethinking Criminal Law*,<sup>22</sup> Fletcher repeats his insistence on maintaining the distinction between the definition of the prohibitory norm and the justifications for violating it.<sup>23</sup> As Fletcher puts it:

Collapsing the distinction between definition and justification eliminates the distinction between conduct that is perfectly legal and conduct that nominally violates a norm but is justified by the assertion of a superior interest or right. It treats killing a human being in self-defense on a par with hunting and killing a coyote. It suggests that a physician's pounding a patient's chest is of the same order as pounding a nail.<sup>24</sup>

Fletcher then draws upon the distinction between definitions and justifications to support the view that the unknowingly justified and hence culpable actor does not merit a justification defense. First, unlike the elements in the definitions of prohibitory norms, the presence of which *mandate* a course of conduct (abstention), justificatory facts merely *privilege* but do not *compel* a course of conduct (violation of the prohibitory norm).<sup>25</sup> The relevance of this point is that not every justified violation of the criminal law can be said to produce a socially preferred state of affairs. From the standpoint of society, it may be a matter of indifference whether the victim of an attack does or does not employ self-defense.

Second, because justifications are exceptions to prohibitory norms, it is proper that one *merit* those privileges in order to invoke them. And the requirement of merit entails the requirement that one be aware of the justifying facts. "For someone who violates a norm to deserve exceptional treatment, he or she must at least know of the circumstances supporting the claim of an exception."<sup>26</sup>

\* \* \*

I want to argue here that both Robinson and Fletcher score points against the other, and that the proper approach to the unknowingly justified actor lies somewhere between the positions.

First, I want to agree with Fletcher with respect to self-defense. I have previously argued that self-defense frequently, and perhaps always, fits the paradigm of excuse better than the paradigm of justification. The latter is represented by the lesser evils defense. If it is socially preferable for V's house to be burned down than to have the forest fire destroy 1,000 houses in the town, burning down V's house is truly a lesser evil and is justified for that reason. In cases of self-defense, however, where the victim may be an innocent attacker—a psychotic, a child, someone who has been duped, and so forth—and may even outnumber the defender—for example, two children with guns, or six dupes who

---

22. George P. Fletcher, *Rethinking Criminal Law* (Little, Brown & Co. 1978).

23. *See id.* at 555-62.

24. *Id.* at 561.

25. *Id.* at 563-64.

26. *Id.* at 565.

have been told the defender is about to detonate an atomic bomb—it is difficult to conclude that it is socially preferable for the attacker or attackers to die instead of the defender. Even the culpable aggressor—who, keep in mind, in cases of self-defense, has not yet finally launched his attack, and is killed or wounded preemptively—possesses only the culpability of one who intends to attack the defender but who might conceivably change his mind if not preemptively attacked.<sup>27</sup> For these reasons, self-defense looks more like a purely personal justification, not a lesser evil, and should function more as excuses function, not justifications.<sup>28</sup> For example, I have argued against allowing non-threatened third parties to aid self-defenders in repelling innocent aggressors, even though were self-defense a justification on a par with lesser evils, third parties should be able so to aid.<sup>29</sup>

So with respect to self-defense at least, I think Fletcher is right to resist extending the defense to the unknowingly justified actor. When a child-hating unknowingly justified “defender” mows down six small children, who unbeknownst to him were about to accidentally kill him by detonating some dynamite they had found, there is reason not to allow him to escape murder convictions on grounds of self-defense.

On a second point, however, my sympathies lie with Robinson rather than Fletcher. When the justification at issue is not self-defense but is instead lesser evils, so that the socially preferable course of action is clear, the unknowingly justified actor looks much more like an attempter than one who has successfully committed a crime.

Consider the following lesser-evil scenario. A dam has burst upriver, and a flood surge is approaching Town. If it reaches Town, it will wreak death and destruction. Fortunately, there is a diversion channel between the head of water and Town. X, who knows of the danger, is attempting to open the valve that would divert the flood into the channel. If X succeeds, Town will be saved, but farmer V’s crops will be flooded and ruined. Clearly, the loss of V’s crops is a lesser evil than death and destruction in Town, and X will have a lesser evils defense to a charge of knowing destruction of V’s property.

Now X is having difficulty opening the valve, so he asks D for instructions on how to do it. D, who hates V, and who is unaware of the flood, thinks X is attempting maliciously to destroy V’s crops and is happy to help. He gives X the assistance sought, X opens the valve, V’s crops are flooded, and Town is saved. X, as stated, has a lesser evils defense with respect to the loss of V’s crops. But what should happen to D, who gave him assistance and knew nothing of the flood?

Fletcher would have to deem D guilty of destroying V’s crops even though he merely aided the justified principal X. For Fletcher, X violated the prohibitory norm, and D’s assistance to X puts D on a par with X in that respect. And

---

27. See Larry Alexander, *A Unified Excuse of Preemptive Self-Protection*, 74 Notre Dame L. Rev. 1475, 1500-01 (1999).

28. *Id.* at 1494-98.

29. *Id.* at 1497.

because D did not know of X's justification, D is guilty, not merely of the crime of attempting to flood V's crops, but of the completed crime of actually doing so.<sup>30</sup>

Robinson, on the other hand, would deem D only an attempter in the flooding of V's crops, not one who is guilty of the completed crime. D's ignorance of the justification renders him culpable, but the net social benefit he produces precludes his being a successful wrongdoer. And I believe Robinson's view of the matter is more persuasive. It really does seem plausible in a case like this to net out the benefits and harms and conclude that, like the typical attempter, D caused no (net) social harm. V's crops are gone, of course, but lives and other property have been saved. The world is better off for D's action, even though D is culpable for having taken it.

But suppose someone in the Fletcher camp were to object to netting out the social costs and benefits this way—that is, summing them across persons, so that V's loss is offset by gains to those in Town. Or rather, they object to netting them out, not in general—after all, they endorse giving the actor who *knows* about the flood a lesser evils defense that is premised on precisely that netting out of costs and benefits—but only with respect to D, who perceived only the costs to V and not the offsetting benefits to Town. In other words, because D himself did not net out the costs and benefits, neither should we in D's case.

Three replies. First, suppose that instead of X attempting to turn the valve and seeking D's assistance, it is farmer V who is doing so. V has learned of the flood and has decided that it is better for him to flood his own farm and suffer a loss than to protect his farm at the expense of death and destruction in Town. Does substituting V for X make a difference?

For Fletcher it does because V is consenting to the destruction of his crops; and Fletcher regards that consent as negating any violation of the prohibitory norm. However, it seems unduly strained to have D's liability for attempt versus success turn on whether it is X or V he is assisting, or on whether V is "consenting" as opposed to "acquiescing" in the destruction of his crops. After all, V is legally *required* to acquiesce if he is aware of the threat to Town; he could not legally resist an obvious attempt to save Town at the cost of his farm. And given that he cannot legally resist, he might decide that he, rather than X, should turn the valve. It seems unplausible, therefore, that D's liability should turn on whether it is X or V whom he aids in turning the valve.

Or consider the case in which P, a member of the Gary Gilmore firing squad and en route to perform his duties in that regard, encounters D and asks him for a bullet, telling him that "I'm going to kill Gary Gilmore today but forgot to load my gun." D is unaware that Gary Gilmore has been caught and tried and is awaiting execution. Rather, D believes P is going to murder Gilmore. Nonetheless, because D hates Gilmore and happens to have a bullet that fits P's rifle, D happily complies with P's request. When P does kill Gilmore, his act is, of

---

30. And of course, the result would be the same for Fletcher were X's and D's roles reversed, and D were the unknowingly justified principal and X the knowingly justified aider. X would have aided D's completed crime but would get the lesser evils defense.



course, fully justified. D, on the other hand, who is complicit in the killing, is no doubt an attempted murderer.<sup>31</sup> But it strikes me as quite implausible that he should be deemed to have murdered Gilmore through the agency of P, the lawful executioner.

Third, consider this variation on the three prospectors hypothetical.<sup>32</sup> D, a radiologist, hates his patient V, who is currently unconscious and undergoing minor surgery. D decides to deliver an intense dose of radiation to V's tongue, which D knows will cause V to develop a fatal cancer at that site in twenty years. Unbeknownst to D, V currently has a cancerous growth at that site, one that will kill him within a year if it is not irradiated. D's action, therefore, rather than shortening V's life by some amount, actually adds nineteen years to it. D has surely attempted to shorten V's life and should be deemed an attempted murderer. However, when V dies in twenty years as a result of D's act, D should not then be deemed to have murdered V. But for D's act, V would have died years earlier. The costs and benefits of D's act net out in its favor; and the netting out of costs and benefits is intra, not inter, personal. They are costs and benefits to V and no one else.

Of course—and this is perhaps the most important point of this analysis of the debate over the unknowingly justified actor—none of these problems of characterization would occur if the criminal law punished attempted and successful crimes the same, as I have argued it should.<sup>33</sup> Almost everyone agrees that the unknowingly justified actor is an attempter. The only question is whether he is a succeder as well. *But nothing should turn on this.* Whether it is X or (the unrecognized) V whom D aids in turning the valve; D's punishment should be the same. Whether D aids a murderer or, unbeknownst to him, an executioner, D's punishment should be the same. Whether D's attempted murder turns out to shorten V's life or lengthen it, the result should be the same.

Mitchell Berman has persuasively argued that the issue of the unknowingly justified actor cannot be settled through conceptual analysis but is instead a question of policy.<sup>34</sup> I believe Berman is correct. I also believe that no valid penal policy supports the attempt/success distinction.

The problem of the unknowingly justified actor is a window on the attempt/success, culpability/wrongdoing, subjectivist/objectivist debate in criminal

31. See e.g. Model Penal Code § 5.01(3) (deeming those who undertake acts that would make them complicit in crimes were those crimes committed guilty of attempts to commit those crimes in the event the crimes are not committed).

32. In the three prospectors hypothetical, prospectors A and B each privately wish to kill prospector C. One night, A secretly pours out the water in C's canteen and replaces it with poison. Later that night, B, unaware of A's act, and believing the liquid in C's canteen is water, pours out the poison and fills the canteen with sand. The next day C goes off in the desert by himself and eventually dies of thirst. B claims, correctly, that had he not emptied C's canteen, C would have died hours sooner of poison. Therefore, B argues that he prolonged C's life and should be guilty only of attempted murder.

33. See e.g. Larry Alexander, *Crime and Culpability*, 5 J. Contemp. Leg. Issues 1, 17-22 (1994).

34. See Mitchell N. Berman, *Justification and Excuse, Law and Morality*, 53 Duke L.J. 1, 58-62 (2003).

2004]

*UNKNOWINGLY JUSTIFIED ACTORS*

859

law. And I believe consideration of the Robinson-Fletcher debate should lead one to favor the attempt-culpability-subjectivist side.

