

PERMISSIBLE PARTS IN IMPERMISSIBLE WHOLES & IMPERMISSIBLE PARTS IN PERMISSIBLE WHOLES: THE CONTRIVED DEFENSE AND DETERRENT THREAT DOCTRINES

[DRAFT— Last Revised 11/30/11]

Russell L. Christopher*

What is the relationship between the permissibility/impermissibility of the part and the permissibility/impermissibility of the whole? Does the moral or legal status of a constituent part of an actor's course of conduct govern the status of the actor's whole course of conduct or, conversely, does the moral and legal status of the actor's whole course of conduct govern the status of the constituent parts? This broader issue is examined in the more specific contexts of the actio libera in causa and deterrent threat doctrines. The latter doctrine maintains that an otherwise impermissible act of carrying out a threatened action may be rendered permissible if embedded within an overall permissible course of action including the issuance of a deterrent threat that fails to induce compliance. The actio libera doctrine addresses the morality and legality of an actor who contrives or culpably causes the conditions of her own defense. This paper considers three principal approaches to the actio libera doctrine. First, hold the actor liable despite granting the defense—the permissibility of the constituent part relating to the defense has no effect on the impermissibility of the actor's whole course of conduct. Second, hold the actor liable by barring the defense—the impermissibility of the actor's whole course of conduct renders the otherwise permissible constituent part relating to the defense also impermissible. Third, the actio libera and deterrent threat doctrines are sufficiently related such that the deterrent threat doctrine supports the second approach to the actio libera doctrine—in both, the permissible/impermissible status of the whole governs the status of the part. This paper challenges each of these principal approaches to the actio libera doctrine.

* Professor of Law, The University of Tulsa College of Law. Thanks to Kathryn Christopher for her criticisms of earlier drafts of this essay.

INTRODUCTION

The *actio libera in causa* doctrine addresses whether actors who satisfy a defense to criminal liability should nonetheless be held liable if they contrived or culpably caused the conditions of their own defense. Such actors satisfy the elements of a criminal offense as well as the requirements for a defense to the offense. Under the doctrine, however, at least some such actors may nonetheless be held liable because of their culpably contriving or causing the conditions or circumstances of their defense. Their culpably contriving or causing the conditions or circumstances of their defense has an inculpatory effect that negates, in some sense, the otherwise exculpatory effect of their defense. Such actors are held just as criminally liable for their commission of an offense as if they had no defense at all.

The intuition supporting the doctrine seemingly has near universal appeal. The doctrine is recognized across a wide number of jurisdictions and legal cultures—both common law and civil law systems--and is applied to a wide variety of different defenses, including both justification and excuses. (Of course, the precise contours of the doctrine may vary across different jurisdictions and different defenses.)

To illustrate the considerable appeal of the doctrine, compare two actors who each commit a crime while substantially intoxicated. Ian innocently becomes substantially intoxicated and commits an offense. Carl contrives to become substantially intoxicated

so that, while in that state, he will commit an offense against his hated enemy but satisfy a defense and not be held liable. Perhaps Ian and Carl may equally satisfy the requirements for a defense based on lack of mens rea or responsibility due to each actor's substantial intoxication. However, under the *actio libera* doctrine only Carl might nonetheless be held liable. His contriving to satisfy the conditions of a defense allows him to be held liable despite otherwise satisfying the requirements of a defense. Even if both actors are equally denied a defense based on lack of mens rea or responsibility, Carl is surely, in some sense, comparatively more culpable and deserving of criminal liability.

Despite the considerable intuitive appeal of the doctrine, problems in the application, scope, and rationale of the doctrine remain. This essay challenges three aspects of the *actio libera* doctrine: (i) the connection between the doctrine and deterrent threats, (ii) a particular version of the doctrine's application to cases of self-defense, and (iii) the limitation on the scope of the doctrine to defenses.

I. CONVERGING THE PARALLEL ISSUES OF, AND COLLAPSING THE DISTINCTION BETWEEN, CONTRIVED DEFENSES AND DETERRENT THREATS

Claire Finkelstein and Leo Katz argue that the deterrent threat doctrine may bear on the contrived defense doctrine.¹ The deterrent threat doctrine addresses the permissibility of (i) threatening to do that which one has no right to do (and would be wrong to do) in order to deter wrongful conduct, and (ii) if the threat fails to deter, carrying out the threatened action that is prima facie wrongful. For example, may one

¹ Claire Finkelstein & Leo Katz, *Contrived Defenses and Deterrent Threats: Two Facets of One Problem*, 5 OHIO ST. J. CRIM. L. 479 (2008).

threaten to use deadly force in defense of property and, if the threat fails to deter, follow through on the threat and use deadly force? There are two principal approaches.

Under what Finkelstein & Katz term the “Backward Induction View,”² the answer is no. The permissibility of the threat depends on the permissibility of the act threatened. Because it is generally impermissible to use deadly force to defend property, it is also impermissible to threaten to use deadly force to defend property.

But under what Finkelstein & Katz term the “Bootstrapping Argument,”³ the answer may be yes. Utterance of the threat may be permissible because it may make the use of any force, even non-deadly force, unnecessary. And carrying out the threatened action, in the event that the threat fails to induce compliance, may also be permissible as a way of making the initial threat credible and thus effective as a deterrent device. The permissibility of both threat and follow through is based on the entire package being an effective deterrent to the use of any force, even non-deadly force. The overall course of conduct—threat and follow through—is permissible to the extent that the interest being protected is sufficiently important and that it will tend to reduce the amount of force employed in the aggregate. That is, the overall societal benefit exceeds the cost. As a result, the overall permissibility of the entire course of conduct renders each prima facie impermissible component as ultimately permissible. The permissibility of the whole governs the permissibility of the part.

² *Id.* at 495.

³ *Id.* at 503.

Finkelstein & Katz ingeniously point out that the two seemingly disparate issues of contrived defenses and deterrent threats are two facets of the same problem. They are logically parallel and symmetrical. The two issues bear a converse relation. The deterrent threat situation involves a prima facie impermissible act embedded in an overall permissible course of conduct (under the Bootstrapping Argument); the contrived defense situation involves a prima facie permissible act embedded in an overall impermissible course of conduct. But in each situation, the agent “causes himself to perform an act whose moral character is at odds with the overall plan from which he acts.”⁴ In general, Finkelstein & Katz argue that the preferable resolution of these issues is that the moral status of the whole or overall plan governs the moral status of a part or particular act of the whole or overall plan.⁵ Thus, regarding contrived defenses, the overall impermissible course of conduct renders what is prima facie permissible (because it otherwise satisfies a defense) as ultimately impermissible. And, regarding deterrent threats, the overall permissible course of conduct renders what is prima facie impermissible as ultimately permissible. The status of the whole controls the status of the part.

⁴ *Id.* at 504.

⁵ For an argument that the status of the whole does not invariably govern the status of the part in the context of punishment, see Russell L. Christopher, *Time and Punishment*, 66 OHIO ST. L.J. 269 (2005) (arguing that the principle of the status of the whole governing the status of the part only applies under consequentialist but not retributive theories of punishment).

This general solution suggests the preferability and consistency of specific solutions to each issue. Joachim Hruschka's secondary duty solution⁶ preferably resolves the contrived defenses issue and the Bootstrapping solution preferably resolves the deterrent threat issue. The two specific solutions are consistent in that under each the status of the whole governs the status of the part. And each specific solution is mutually supporting. If one is right the other must be right; if one is wrong the other must be wrong. As a result, if we agree with one specific solution, the cost of disagreeing with the second solution is disagreement with the first solution.

Let us attempt to formalize this relationship:

Contrived Defense: Actor commits prima facie permissible act X in an overall impermissible course of conduct. Under Finkelstein & Katz's approach, the overall impermissibility of the course of conduct governs the status of the prima facie permissible act X. Thus, the prima facie permissible act X is rendered impermissible. A prima facie permissible act—because it otherwise satisfies a defense—is rendered

⁶ Finkelstein & Katz explain Hruschka's account of the contrived defenses issue as follows:

[F]undamentally Hruschka's approach has us evaluate the defendant's conduct in its totality, and then assess the individual components of that course of action accordingly. Hruschka has us look at the course of conduct as a whole, and if we find that we disapprove, to recognize a secondary duty [in addition to the primary duty not to violate statutory norms] not to create situations that would require the defendant to perform the illegal, but ostensibly justified or excused, action.

Finkelstein & Katz, *supra* note 1, at 504. For Hruschka's own explanation of his approach, see *id.* at 480 n.2 (citing JOACHIM HRUSCHKA, STRAFRECHT NACH LOGISCH-ANALYTISCHER METHODE (1983)).

impermissible if it is a part of an overall impermissible plan of contriving to satisfy a defense.

Deterrent Threat: Actor commits prima facie impermissible act Y in an overall permissible course of conduct. Under Finkelstein & Katz's approach, the overall permissibility of the course of conduct governs the status of the prima facie impermissible act Y. Thus, the prima facie impermissible act Y is rendered permissible. A prima facie impermissible act of carrying out a threatened action is rendered permissible if part of an overall permissible plan of deterring wrongful conduct.

As a result, the seemingly disparate problems of contrived defenses and deterrent threats share a common solution. The status of the whole governs the part. The seemingly disparate problems are the reverse of each other, two facets of each other, or are parallel to each other. The single solution yields opposite outcomes for each type of problem. The overall impermissibility of the plan in a contrived defense situation renders the prima facie permissible part satisfying a defense impermissible; the overall permissibility of the plan in a deterrent threat situation renders the prima facie impermissible part of carrying out the threatened action permissible.

Note that the premise of Finkelstein & Katz's approach is that the issues of contrived defenses and deterrent threats are reverse or mirror images of each other; two distinct and opposite facets of the same problem. That is, that the prima facie permissible act satisfying a defense (ultimately deemed impermissible) and that the prima facie impermissible act of carrying out the threatened action (ultimately deemed permissible)

are distinct. With respect to the above formalizations, acts X and Y must be distinct.

Otherwise, if the distinct acts of X and Y were actually one and the same act, Finkelstein & Katz's solution would yield contradictory results. The single act would be ultimately impermissible under the contrived defense approach and ultimately permissible under the deterrent threat approach.

May the distinct acts be converted into a single, identical act? That is, can a single act be both prima facie permissible as satisfying a defense (but ultimately impermissible because part of a contrived defense) and prima facie impermissible as carrying out a threatened action (but ultimately permissible as part of an overall permissible deterrent threat plan)? This essay argues that the claimed distinct and parallel doctrines of contrived defense and deterrent threat may converge and collapse into each other.

Consider the hypothetical that Finkelstein & Katz utilize as an example of the deterrent threat doctrine.⁷ A informs B, "If you steal my property I will kill you." The deterrent threat is issued with either the intent or knowledge or awareness that it will (likely) deter B. After B steals A's property, A kills B. Under the deterrent threat doctrine, Finkelstein & Katz argue that the prima facie impermissible killing of B is permissible because it is part of a permissible overall course of conduct. Carrying out the otherwise impermissible threatened action is rendered permissible following the issuance of a deterrent threat that fails to deter. Because the overall course of conduct is permissible, the individual act of killing is also permissible.

⁷ Finkelstein & Katz, *supra* note 1, at 481.

But consider the following supplement to the original hypothetical—Variation 1—that seeks to combine the seemingly distinct deterrent threat and contrived defense situations into a single situation. Suppose that *A* issues that same threat not only with the intent or knowledge or awareness that it will (likely) deter, but also with the intent or knowledge or awareness that *B* may be provoked into stealing the property and that *A*'s killing of *B* will enjoy a defense of defense of property. Is *A*'s killing of *B* permissible or impermissible? Under the deterrent threat doctrine approach favored by Finkelstein & Katz, the killing is permissible as part of an overall permissible plan. But under the contrived defense approach favored by Finkelstein & Katz, the killing is impermissible as part of an overall impermissible plan. As a result, whether the killing is permissible or impermissible seems indeterminate under Finkelstein & Katz's approach.

To this claim of indeterminacy, perhaps one might object that Variation 1 fails to qualify as a true contrived defense situation. The killing is not even *prima facie* permissible because it fails to satisfy a defense. Generally, killing to prevent theft of property is not conduct that satisfies a defense. As a result, the claimed indeterminacy is merely the result of a situation that fails to satisfy both the contrived defense and deterrent threat paradigms. As such, the claim of indeterminacy fails.

Even if this is not a true contrived defense situation, however, the claim of indeterminacy may still stand. Moreover, the claim may even be strengthened. The killing is not only impermissible from a contrived defense perspective but also from a “conventional” defense perspective—it is doubly impermissible, so to speak. But from

the deterrent threat perspective it is permissible. So, is the killing permissible or impermissible? Again, the answer seems indeterminate.

In any event, let us consider another hypothetical—Variation 2—that perhaps avoids the above objection by truly qualifying as a contrived defense situation. Suppose again *A* informs *B*, “I will kill you if you steal my property.” *A* issues this threat with the awareness that it will (likely) deter *B* from stealing but with the intent that it will provoke *B* into trying to kill *A* whereby *A* will kill *B* in self-defense.⁸ *B* steals the property and attempts to kill *A*. *A* kills *B* and asserts the defense of self-defense. Is the killing permissible or impermissible? From a contrived defense perspective, the prima facie permissible killing in self-defense is rendered impermissible because it is a contrived defense and thus part of an overall impermissible plan. But from a deterrent threat perspective, the prima facie impermissible act of killing⁹ is rendered permissible because it is part of an overall permissible plan of carrying out a threatened act after the issuance of a deterrent threat failed to deter. Again, whether the killing is permissible or impermissible under Finkelstein & Katz’s approach would seem to be indeterminate.

One might object that while this situation (Variation 2) may conform to the contrived defense paradigm, it does not truly conform to the deterrent threat paradigm. The act of killing was not merely the carrying out of a threatened action that is prima facie impermissible—killing to prevent property theft. It was also a defensive killing in

⁸ Alternatively, the scenario could be the opposite. *A* issues the threat with the intent that it will deter and the awareness that it will likely provoke *B*.

⁹ The act of killing is prima facie impermissible either because it is carrying out an impermissible threatened action or because it is a contrived defense.

response to *B*'s attempt to kill *A*. While true, this may not matter. Either way the killing is impermissible in isolation from *A*'s overall permissible deterrent threat plan. Moreover, this only makes the killing doubly impermissible, so to speak. It is impermissible not only as carrying out a threatened action that is prima facie impermissible—killing to prevent the theft of property—but also impermissible as part of a contrived defense. All that the deterrent threat paradigm requires is a prima facie impermissible act—whatever the source of the impermissibility—embedded within an overall permissible deterrent threat plan.

In any event, let us consider yet another hypothetical—Variation 3—that attempts to avoid the objection. Variation 3 features an actor contriving a defense neither on defense of property as in Variation 1 nor on self-defense as in Variation 2. Rather than a traditional defense, the actor contrives a defense based on the deterrent threat doctrine itself. Suppose *A* again informs *B*, “I will kill you if you steal my property.” *A* issues the threat with the awareness that it will (likely) deter *B* from stealing the property but with the intent that if *B* is not deterred *A* may permissibly kill *B* in reliance on the deterrent threat doctrine. That is, *A* is both issuing a deterrent threat and contriving a (moral) defense based on the deterrent threat doctrine. Suppose *B* attempts to steal the property and *A* kills *B* to prevent the theft.

Is the killing permissible or impermissible? From a contrived defense perspective, the killing is (i) impermissible because it fails to satisfy a criminal defense, and (ii) prima facie permissible because it is part of an overall permissible

plan vis a vis the deterrent threat doctrine but (iii) ultimately impermissible as part of a contrived (moral) defense. From a deterrent threat perspective, carrying out the threatened action is (i) prima facie impermissible—both because it fails to satisfy a defense and because it is a contrived defense—but (ii) ultimately permissible because it is part of an overall permissible plan of carrying out a threatened act after the threat fails to deter. Thus while the contrived defense perspective finds it impermissible, the deterrent threat approach finds it permissible. Again, under Finkelstein & Katz’s approach, whether the killing is permissible or impermissible seems indeterminate.

In the vast majority of cases, the link between the two seemingly disparate doctrines that Finkelstein & Katz have ingeniously forged may well hold. But the above hypotheticals suggest that there may be some idiosyncratic cases where an actor simultaneously makes a deterrent threat and contrives a defense. In such cases, the general solution offered to resolve both the deterrent threats and contriving defenses issues--the status of the whole governs the status of the individual part--fails. The actor’s overall plan and its constituent parts would be viewed as permissible under (the Bootstrapping Argument approach to) the deterrent threat doctrine but impermissible under (some approaches to) the contriving defenses doctrine. The resulting indeterminacy of an actor’s overall plan and its constituent parts being both permissible and impermissible may undermine any conceptual link between the two doctrines.

However, one approach to the contriving defense issue is only strengthened by the above indeterminacy. Paul Robinson's approach does not incorporate the principal that the status of the whole governs the status of the parts. Unlike other approaches to the contriving defense issue, Robinson grants the defense to the contriving actor while nonetheless holding the actor liable. Thus, the impermissible overall plan of conduct of the defense contriver does not govern the status of all of the individual parts.

II. ALLOW THE DEFENSE BUT STILL HOLD THE ACTOR LIABLE—ROBINSON'S APPROACH

Paul Robinson's approach to the *actio libera* doctrine differs from the traditional and predominant approach.¹⁰ Under the traditional/predominant approach, an actor who contrives a defense is barred from successfully obtaining the defense. The contriving of the defense invalidates the defense. Robinson argues that the preferable approach is to allow the defense but nonetheless hold the actor liable for the offense. Though an actor may have a defense for the last act undertaken, the actor may still be liable for the very conduct, if undertaken culpably, that earlier constituted the contriving of the defense. To illustrate this, Robinson supposes that an actor wishes to burn down an enemy's house but avoid criminal liability for arson.¹¹ The actor intentionally/culpably sets a fire in a forest that threatens a nearby town with the intent/culpability that the enemy's house would be destroyed

¹⁰ Paul H. Robinson, *Causing the Conditions of One's Own Defense: A Study in the Limits of Theory in Criminal Law Doctrine*, 71 VA. L. REV. 1 (1985).

¹¹ *Id.* at 28.

in a firebreak necessary to save the town. The actor asserts a necessity or lesser evils justification for setting the firebreak that destroyed the enemy's house but saved thousands of the endangered townspeople's lives. Under the traditional/predominant approach the actor would be denied the lesser evils defense and would be held liable for arson of the enemy's house. By contriving the defense, the actor loses the defense.

Robinson argues that despite the actor's bad motive in culpably causing the defense, society still benefits if the culpable actor sets the firebreak and saves the townspeople's lives. Rather than bar the defense and discourage beneficial conduct, such conduct is affirmatively good and needs to be encouraged. The actor's conduct in setting the firebreak is thus justified and the earlier conduct of contriving the defense does not invalidate the defense. Nonetheless, under Robinson's approach, the actor may still be held liable for intentionally destroying the enemy's house. Setting the initial fire is the actus reus and the but for cause of the destruction of the enemy's house. That actus reus was also accompanied by a culpable mental state toward the destruction of the enemy's house. So, based not on the later act of setting the firebreak but on the earlier act of setting the forest fire that necessitated the firebreak, there is the actus reus, mens rea, and causation of the crime of arson for destruction of the enemy's house.

The claimed advantage of Robinson's approach, however, is not yet clear. If the actor will be held liable for the offense of arson of the enemy's house, what incentive does the actor have for setting the firebreak? Avoiding the additional

offense of homicide counts regarding the deaths of the townspeople. But the actor has this incentive under the traditional/predominant approach as well. Even without granting the actor a lesser evils justification for the arson offense, the actor still has an incentive not to add to his liability the deaths of thousands of townspeople. What incentive does the granting of a justification defense create over and above the incentives already in place under the traditional/predominant approach? Perhaps little.¹²

The clearer advantage of Robinson's approach relates less to the actor than to the rights of other parties who may become involved. First, consider the possible involvement of the property owner. Under the traditional/predominant approach that bars a lesser evils justification to the actor in setting a firebreak, the actor's conduct would be unjustified thereby allowing the homeowner to justifiably prevent the firebreak necessary to save thousands of lives. Because there is a societal benefit in setting the firebreak, the property owner should not be able to justifiably resist the setting of the firebreak. The traditional/predominant approach generates an unfortunate outcome. In contrast, Robinson's approach generates a preferable outcome. By granting a lesser evils justification to the actor for setting the firebreak, the actor is acting justifiably and the property owner could not justifiably prevent

¹² Robinson maintains that "the unavailability of the defense may reduce his incentive to set the firebreak and save the town." *Id.* But Robinson may acknowledge that the actor has an incentive to set the firebreak even without granting a justification defense to the actor. *See id.* (noting that "the actor retains a special incentive to set the firebreak, for if the town burns down or injures someone, he may be liable for this additional harm").

the actor from setting the firebreak. The actor's justified conduct could not justifiably be resisted.

Second, Robinson's approach generates a preferable outcome with respect to a third party who might aid the actor in setting the firebreak. Under the traditional/predominant approach, a conceptual anomaly arises. The actor would be unjustified (as a contriver) but the third party would be justified for performing the same conduct. Under Robinson's approach, both the actor and the third party would be justified thereby avoiding the anomaly.

Curiously Robinson fails to mention what might be the even stronger argument in favor of his approach and against the traditional/predominant approach. If under the traditional/predominant approach, the actor is unjustified in setting the firebreak a third party might be precluded from aiding the actor set the firebreak necessary to save thousands of lives.¹³ In aiding the setting of the firebreak, the third party would be aiding unjustified conduct. Generally, aiding unjustified conduct is itself unjustified. Robinson's approach avoids this unfortunate outcome by granting a justification to the actor. Thus, a third party may justifiably aid the justified conduct of the actor. (Justified conduct may justifiably be aided.) In any event, Robinson's approach does yield some advantages over the traditional/

¹³ Robinson seems to assume that under the traditional/predominant approach a third party would be justified in aiding the setting of the firebreak despite that the actor would be unjustified in doing so. *See id.* Thus, an actor may justifiably aid unjustified conduct. But in analyzing the rights of the property owner under the traditional/predominant approach, Robinson assumes that the unjustified conduct of the actor could be justifiably resisted. If only unjustified conduct may be justifiably resisted, why can unjustified conduct be justifiably aided?

predominant approach with respect to the rights of other persons in lesser evils cases.

But Robinson's approach may not work as well for self-defense cases as it does for necessity cases. I will argue that it does not and then sketch some reasons for the possible discrepancy between its greater success with lesser evil cases and its lesser success with self-defense cases.

Consider the following hypothetical. *A* tricks *B* into believing that *A* is about to imminently and unlawfully kill *B*. *A* does so with the intent that (i) *B* will defend herself by trying to shoot and kill *A*, (ii) *A* will thereafter shoot and kill *B*, and (iii) *A* will escape criminal liability by claiming justified self-defense. Perhaps *A* uses a fake gun to trick *B*. As *B* is drawing a real gun to shoot *A*, *A* more quickly draws a real gun and kills *B*. *A* claims the justification defense of self-defense. Analyzed under the traditional/predominant approach, *A*'s contriving the defense invalidates the defense and *A* would be held liable for the homicide of *B*. Under Robinson's approach, *A*'s act of shooting and killing *B* is justified in self-defense. Despite contriving the defense, *A* may still successfully obtain the defense. But *A* is still held liable for *B*'s murder. *A*'s earlier act of tricking *B* is the actus reus, accompanied by a culpable mens rea of intent to kill *B*. And *A*'s earlier act of tricking *B* into trying to kill *A* is the but for cause of *B*'s death. Under both approaches, *A* will be criminally liable for *B*'s death but only under Robinson's approach will *A* be justified for the act of shooting and killing *B*.

But consider the involvement of a third party. Suppose a third party arrives on the scene with perfect knowledge of the situation. Suppose further that the third party's only choices are to kill *A*, kill *B*, or do nothing. What should the third party do? Under Robinson's approach, *A*'s act of shooting *B* is justified so the third party must not interfere with *A*'s justified act of shooting *B*. If the third party intervenes and shoots anyone it must be *B* who is acting unjustifiably in shooting at *A* when *A* does not (yet) pose an actual threat to *B*. Though Robinson does not explicitly apply his alternative approach to the *actio libera* doctrine to examples involving self-defense, I assume that this is what his approach would yield. In contrast, under the traditional/predominant approach, it is *A* who is unjustified in shooting *B*. *A*'s contriving to kill *B* in self-defense invalidates *A*'s justification. Thus *A* is unjustified. As a result, under the traditional/predominant approach, the third party would be justified in shooting and killing *A*—the unjustified actor.

In this example, the traditional/predominant approach and Robinson's approach yield opposite outcomes as to the rights of a third party. According to Robinson's approach the third party is only justified to intervene and kill *B*; under the traditional/predominant approach, the third party may only intervene and justifiably kill *A*. Unlike in the lesser evils example above, Robinson's approach seems counter-intuitive and wrong as applied to the self-defense case. If the third party must choose who should die between *A* and *B*, *A* should be the one to die. *B* is the innocent party who was tricked by *A*. *A* is the party in the wrong. Even under Robinson's approach, *A* is wrong for having unlawfully and unjustifiably killed *B* (by *A*'s earlier act of contriving a defense). So, then, why should, under Robinson's

approach, the third party be obliged to side with the wrongful party, *A*, who will be criminally liable for unjustifiably killing *B*?

Our intuitions about the lesser evils example and self-defense example diverge. In the lesser evils example, we think that a third party should aid the contriving actor commit the act for which she has contrived a defense. But in the self-defense example, it is at the very least less clear that the third party should aid the contriving actor commit the act for which she has contrived the defense. Moreover, in the self-defense example, not only is it less clear that the third party should aid the contriving actor, but also the third party should oppose and prevent the actor from committing the act for which she has contrived a defense. While it is clear in the lesser evils example that society benefits from either the contriving actor or the third party or both setting the firebreak, it is not clear that society benefits from *A* killing *B*. Moreover, societal interest would seemingly be aligned with opposing the actor (*A*) who has wrongfully contrived a defense (and who will be liable for homicide) and protecting the innocent actor (*B*).

While Robinson's approach may yield the better outcome in the lesser evils examples of contrived defenses, the traditional/predominant approach seems to yield the better outcome in the self-defense example.

III. CONTRIVING DEFENSES V. CONTRIVING TO AVOID COMMITTING OFFENSES

The apparently unchallenged and undefended assumption of the *actio libera* or culpably causing one's defense doctrine is that there is something special about contriving a defense. Contriving to avoid satisfying the elements of an offense is different. An actor who successfully contrives to satisfy a defense is nonetheless held liable, but the actor who successfully contrives to avoid satisfying the elements of an offense is not. Both equally contrive but only one is held liable. Why the difference?¹⁴

For example, suppose actor *O* contrives to kill her hated enemy by giving her a plane ticket with the knowledge that statistically 1 in every 500,000 flights crash and with the desire that this flight will crash and that her hated enemy will die from the crash. Or the actor gives the enemy free super-powered motorcycle. Or the actor purchases for her hated enemy a lifetime pass for unlimited access to enjoy Swimming with the Sharks or sky-diving. The actor is contriving to kill her enemy without satisfaction of all the elements of a homicide offense. Suppose the enemy does die from one of these comparatively risky activities. But the actor perhaps fails to satisfy all of the elements of a homicide offense—the actor may lack sufficient mens rea or may not be the proximate cause. Despite contriving to kill her hated enemy, the doctrine of *actio libera* would not hold actor *O* liable. Compare actor *O* with actor *D*, who contrives to kill her hated enemy under the circumstance of some defense, for example, self-defense. Here, the *actio libera* doctrine would hold actor *D* liable for homicide. If each actor equally contrived to kill her hated enemy in such a

¹⁴ For discussion of the ill-defined boundaries between offenses and defenses, see generally Claire Finkelstein, *Positivism and the Notion of an Offense*, 88 CAL. L. REV. 335 (2000).

way as to avoid liability, why is actor *O* not liable and actor *D* is liable? Why is contriving about satisfaction of a defense impermissible but contriving about the non-satisfaction of an element of an offense permissible?

There are a number of possible answers, none of which are entirely satisfactory.

A. *Wrongfulness/Unlawfulness v. Non-Wrongful/Lawful*

One possible answer is that the difference can be explained by the status of each type of conduct the contriving actor contrives to commit. By contriving to commit conduct that satisfies the elements of a criminal offense and satisfies a defense, the actor is contriving to commit wrongful and unlawful conduct. Conduct that satisfies the elements of a criminal offense is wrongful and unlawful. Contriving to commit wrongful and unlawful conduct that also satisfies a defense is nonetheless impermissible and should be treated as wrongful and unlawful. In contrast, contriving to commit conduct that fails to satisfy the elements of a criminal offense is contriving to commit lawful and non-wrongful conduct. Contriving to commit non-wrongful and lawful conduct should, obviously, be permissible and should remain decriminalized. That the actor contrived to commit lawful conduct, rather than accidentally committed lawful conduct, should not be held against the actor. That the actor contrived to commit lawful conduct is not a reason to criminalize his conduct. Moreover, that the actor contrived to commit, rather than accidentally committed, lawful conduct may make the actor even less blameworthy.

But this possible answer is not entirely persuasive. Not all contriving to commit conduct that both satisfies the elements of an offense as well as satisfies a defense is wrongful and unlawful. True, contriving to commit conduct that satisfies both an offense and an excuse defense is both wrongful and unlawful. Excused conduct is both wrongful and unlawful. Excused conduct ought not have been done; it is impermissible. In contrast, justified conduct is neither wrongful and nor unlawful. Justified conduct is permissible. Under some accounts, justified conduct is an affirmative good that ought to be done. In any event, unlike excused conduct, justified conduct is permissible. It is neither wrongful nor unlawful.

Thus this possible answer may succeed but only with respect to excuse defenses. It fails with respect to justification defenses. Both contriving to commit conduct that fails to satisfy the elements of an offense and contriving to commit conduct that satisfies the elements of an offense but also satisfies the elements of a justification defense are each permissible. Each actor is committing lawful and non-wrongful conduct. Despite this similarity, the former contriving is permissible while the latter contriving is impermissible. So again, what is the difference between our contriving rules regarding defenses (including justifications) and offenses?

B. Harm

Another possible answer is that conduct that satisfies both an offense and a defense nonetheless entails a harm but conduct that fails to satisfy an offense fails to constitute a harm. But this does not seem correct with all examples. Consider the example above re actors *O* and *D* killing their respective hated enemies. According

to this possible answer, *D*'s killing of her hated enemy constitutes a harm but *O*'s killing of her hated enemy fails to constitute a harm. Contriving to kill and avoid liability by satisfying a defense constitutes a harm but contriving to kill and avoid liability by non-satisfaction of all elements of an offense fails to constitute a harm.

But this seems wrong. In each case there is a harm—the death of the victim. If the harm of the killing of *D*'s enemy is that a person has died, then that same harm also occurred with respect to the killing of *O*'s enemy—a person has died.

Moreover, if there was any difference between the killings, the *actio libera* doctrine and the harm rationale have it backwards. Whereas the claimed *harmless* killing of *O*'s enemy involved a victim that was a completely innocent person, the claimed *harmful* killing of *D*'s enemy involved in some sense a wrongful aggressor. Under the *actio libera* doctrine and the posited harm rationale, *O*'s killing of a completely innocent person is harmless and entails no liability for *O*. But *D*'s killing of (in some sense) a wrongful aggressor is harmful and entails criminal liability for *D*. Contrary to the outcome under the *actio libera* doctrine and the harm rationale, any difference between the killings should be that the killing of the completely innocent person is comparatively more harmful and a stronger basis for liability.

C. *Legality Principle*

Another possible answer is the Legality principle. To some extent the Legality principle applies to elements of offenses more so than to defenses. As a result, while the Legality principle might block prosecution of an actor who contrives to commit conduct that fails to satisfy the elements of an offense, it may

not block prosecution of an actor who contrives to commit conduct that satisfies both an offense and a defense. Holding an actor liable despite her satisfaction of a defense that she contrived to satisfy perhaps does not violate the Legality principle. But holding an actor liable despite his non-satisfaction of the elements of an offense he contrived to avoid committing would violate the Legality principle.

But the Legality principle is not entirely persuasive as an explanation. The Legality principle may be unable to supply an explanation for offenses containing offense elements that pertain to a defense. Three different categories of offenses containing offense elements pertaining to a defense will be considered: (i) offenses that in some jurisdictions or under some formulations contain offense elements that are explicitly negative elements of defenses, (ii) offenses that invariably and explicitly contain offense elements pertaining to defenses, and (iii) offenses that in some jurisdictions or under some formulations contain offense elements that are implicitly negative elements of defenses.

1. *Offenses that Sometimes Explicitly Contain Negative Elements of Defenses*

Some jurisdictions' homicide and murder offenses contain the offense element of "unlawfully" or "unjustifiably" killing.¹⁵ Thus, a killing that satisfies a justification defense would fail to satisfy this offense element. Suppose, in such a jurisdiction, an actor contrives to kill her hated enemy by tricking or provoking the

¹⁵ See, e.g., CAL. PENAL CODE §187(a) ("Murder is the *unlawful* killing of a human being . . .") (emphasis added); PA. CONS. STAT. § 2503(a) ("A person who kills an individual *without lawful justification* commits voluntary manslaughter . . .") (emphasis added).

hated enemy into aggressing against the actor whereby the actor can justifiably or lawfully employ self-defense. The actor has contrived or culpably caused the killing of another under conditions of lawful and justifiable self-defense. Is contriving to kill under the condition or circumstance of lawful and justifiable self-defense subject to the application of the *actio libera* doctrine and thus the actor may be held liable? Or is it to be understood as contriving to kill by avoiding satisfaction of all of the elements of an offense, and thus not subject to the application of the *actio libera* doctrine, and thus the actor is not liable?

If defenses contained in offense elements are subject to the application of the *actio libera* doctrine, then the Legality principle fails to provide an explanation for the exclusion of contriving to avoid committing offenses from the application of the *actio libera* doctrine. If the *actio libera* doctrine still applies, despite the offense containing the element of “unjustifiably,” and the actor may be held liable for the murder, then the Legality principle fails to block liability despite an actor’s non-satisfaction of the elements of the offense. And thereby, the Legality principle fails to explain why contrivers of non-satisfaction of the elements of offenses should not also be held liable. That is, if one can be held liable despite the non-satisfaction of some elements (those elements pertaining to defenses) then one could also be held liable despite the non-satisfaction of other elements (those elements not pertaining to defenses). As a result, contrivers of the non-satisfaction of the elements of an offense could also be held liable under a broadened *actio libera* doctrine without violating the Legality principle. If neither the Legality principle nor any other

rationale can explain the existing limitation on the scope of the *actio libera* doctrine to contrivers of defenses, then the limitation is arbitrary.

If defenses contained in offense elements are not subject to the application of the *actio libera* doctrine, then the Legality principle might provide an explanation but the existing scope of the *actio libera* doctrine must be narrowed. Rather than applying to any and all defenses, the doctrine would be limited to those defenses that are not contained in offense elements. Such a narrowed *actio libera* doctrine would thus be dependent on whether a defense is exclusively treated as a defense or is contained in an offense element. But many argue that there is little substantive or moral reason for whether a defense is treated exclusively as a defense or as a negative element of an offense.¹⁶ And thus there would be little moral or substantive reason behind such a narrowed *actio libera* doctrine's limitation to defenses that are exclusively treated as defenses. And thus even a narrowed *actio libera* doctrine would be arbitrary.

¹⁶ See, e.g., MODEL PENAL CODE § 1.12, cmt. 4, at 196 (Official Code and Revised Comments 1985) (“There is admittedly no certain principle by which to gauge when a qualification of the scope of a prohibition should be viewed as a matter of excuse or justification as distinguished from an aspect of the basic definition of the crime.”); DOUGLAS HUSAK, PHILOSOPHY OF CRIMINAL LAW 209 (1987) (noting that most theorists find satisfactorily accounting for the distinction to be an “impossibility”); MICHAEL MOORE, ACT AND CRIME: THE PHILOSOPHY OF ACTION AND ITS IMPLICATIONS FOR THE CRIMINAL LAW 178-79 (1993) (opining that the distinction “makes no sense to me”); Paul Robinson, *Rules of Conduct and Principles of Adjudication*, 57 U. CHI. L. REV. 729, 757 n.65 (1990) (“Defenses, even general defenses of justification and excuse, similarly might be defined as negative elements of offenses. Given its manipulable nature, some additional refinement is needed before we can assume that the conceptual structure of the criminal law ought to revolve around the inculpation-exculpation distinction . . .”).

To illustrate this arbitrariness of a narrowed *actio libera* doctrine, consider the following hypothetical. Suppose that an actor contrives to kill two different victims in two different jurisdictions under contrived conditions of lawful, justifiable self-defense. The actor kills *V1* in jurisdiction O that includes in its murder offense the offense element of an unlawful or unjustifiable killing. The actor kills *V2* in jurisdiction D that does not include that offense element in its murder offense. If the *actio libera* doctrine does not apply to defenses when contained in offense elements, then the actor is not liable for murder (or any other offense) in jurisdiction O but is liable for murder in jurisdiction D. Despite equally contriving to commit, and actually committing, the same conduct in each jurisdiction to equally situated victims, the actor is liable for murder in one jurisdiction but liable for nothing in the other.

Either way, the *actio libera* doctrine is arbitrary. Regardless of whether or not the *actio libera* doctrine applies to defenses contained in offense elements, the *actio libera* doctrine is arbitrary. If it does apply, then the limitation on the scope of the doctrine to defenses is arbitrary. If it does not apply, then the doctrine itself is dependent on a distinction between offenses and defenses that many commentators find arbitrary.

This arbitrariness could be avoided in either of two ways. First, the doctrine could apply to both contrivers of defenses as well as contrivers of the non-satisfaction of offenses. Second, the doctrine could apply to neither. Because neither way of avoiding the arbitrariness seems palatable, perhaps there is some other way

of explaining the doctrine's limitation other than by recourse to the Legality principle (and other than by recourse to the wrongfulness/unlawfulness rationale or the harm rationale discussed above). But such an explanation seems elusive.

2. *Offenses that Invariably and Explicitly Contain Elements Pertaining to a Defense*

The previous section considered offenses that only under some formulations or under some jurisdictions contain offense elements pertaining to defenses. Yet another category of offenses containing offense elements pertaining to defenses are those offenses that necessarily or invariably contain offense elements pertaining to defenses. For example, consider voluntary manslaughter under the common law or the extreme mental or emotional disturbance (EMED) variation of Manslaughter under the Model Penal Code. The informal defense of provocation or EMED is always a formal element of the offense.¹⁷ The defense of provocation is also a negative element of the offense of murder.

That the defense of provocation/EMED is both a formal (positive) element of the offense of voluntary manslaughter and a negative element of the offense of murder complicates our analysis. Suppose an actor contrives to kill under circumstances of provocation/EMED. Let us first consider the effect of the contriving on the offense of voluntary manslaughter. The actor is both contriving to

¹⁷ See, e.g., PA. CONS. STAT. § 2503(a) (“A person who kills an individual without lawful justification commits voluntary manslaughter if at the time of the killing he is acting under a sudden and intense passion resulting from serious provocation . . .”); see MODEL PENAL CODE § 210.3(1)(b) (Official Code and Revised Comments 1985) (defining manslaughter, in part, as “a homicide which would otherwise be murder is committed under the influence of extreme mental or emotional disturbance . . .”).

obtain the defense of provocation/EMED as well as contriving to satisfy the offense of voluntary manslaughter/manslaughter. The problem of arbitrariness in the above section does not seem to arise here because the actor is not contriving to avoid satisfaction of the elements of the offense of voluntary manslaughter/manslaughter.¹⁸

But with respect to the offense of murder, the problem of arbitrariness would arise. If one contrives to kill under provocation or EMED, does the *actio libera* doctrine apply and thereby invalidate the exculpatory effect of the mitigation defense leaving the actor liable for murder? Or does the *actio libera* doctrine not apply because the actor is contriving to avoid satisfaction of a negative offense element of murder? If the *actio libera* doctrine does apply, then the actor can be held liable for murder despite non-satisfaction of the elements of the offense, the Legality principle fails to explain why the doctrine is limited to defenses, and this limitation is arbitrary. However, if the *actio libera* doctrine does not apply, then the Legality principle might provide an explanation for the limitation, but the doctrine's application is dependent on what many conclude to be an arbitrary distinction between offenses and defenses. Either way, the *actio libera* doctrine is arbitrary.

¹⁸ But one curious wrinkle does arise if we suppose that the prosecutor only charges the actor with the offense of voluntary manslaughter/manslaughter. And this curious wrinkle may lend some support to the preferability of Robinson's approach over the traditional/predominant approach. Under the traditional/predominant approach to the *actio libera* doctrine, by contriving to satisfy the defense of provocation/EMED, the actor does not satisfy the defense. But if the actor does not satisfy the defense, then the actor also does not satisfy the positive offense element of killing under provocation/EMED. By not satisfying a requisite element of the offense, the actor cannot be held liable for the offense. In contrast, under Robinson's approach, the actor would satisfy the defense, thereby satisfying the positive element of the offense and the actor could be held liable for the offense.

3. *Offenses that Sometimes Implicitly Contain Negative Elements of Defenses*

As an example of this category, consider the defense of renunciation or abandonment to an attempt offense. The defense of renunciation/abandonment is both a formal defense (in some jurisdictions) and an *implicit* negative element of the offense of attempt. (This makes it distinct from a justification defense (e.g., self-defense) that is both a formal defense and an *explicit* negative element of a murder offense.)¹⁹ Abandonment/renunciation serves as an implicit negative element of an offense by negating the mens rea of intent.²⁰ If the actor renounces/abandons, then perhaps the actor never had the sufficiently firm or resolute intention required for attempt liability.²¹

Application of the *actio libera* doctrine to the defense of renunciation/abandonment triggers the same problem of arbitrariness discussed above. Suppose an actor contrives to avoid liability for the commission of an offense by contriving to

¹⁹ PAUL ROBINSON & MICHAEL CAHILL, CRIMINAL LAW 30 (2d ed. 2012) (noting that renunciation is an offense-modification defense to inchoate offenses that is “really just helping to refine the definition of these inchoate offenses”); *id.* at 31 (“An offense modification defense can as easily be drafted as a negative element of the offense itself.”) (emphasis omitted).

²⁰ MODEL PENAL CODE § 5.01 cmt. 8 at 358 (“[E]ven where voluntary desistance was not a defense, abandonment by the actor may have resulted in exoneration by negating a criminal intent.”).

²¹ GEORGE FLETCHER, RETHINKING CRIMINAL LAW 194 (1978) (“Abandonment is accepted as a defense because it reveals that the actor’s criminal will was insufficiently resolute to carry through.”); ROBINSON & CAHILL, *supra* note 19, at 521 (“[A] desert-based rationale justifies a renunciation defense, however, only in cases where the actor’s sole reason for renouncing is that his intention is not (and quite possibly never was) sufficiently resolute.”).

obtain a renunciation/abandonment defense. Does the *actio libera* doctrine apply thereby invalidating the exculpatory effect of the defense and the actor is liable for an attempt? Or does the *actio libera* doctrine not apply because the actor is contriving to avoid satisfaction of a negative offense element and the actor is not liable? If the *actio libera* doctrine does apply, then the actor can be held liable for an attempt despite non-satisfaction of the elements of the offense, the Legality principle fails to explain why the doctrine is limited to defenses, and this limitation is arbitrary. However, if the *actio libera* doctrine does not apply, then the Legality principle might provide an explanation for the limitation, but the doctrine's application is dependent on what many conclude to be an arbitrary distinction between offenses and defenses. Either way, the *actio libera* doctrine is arbitrary.

Application of the *actio libera* doctrine to the renunciation/abandonment defense may pose a special, particularly puzzling challenge. The general precept of the *actio libera* doctrine is that contriving to obtain exculpation through a defense inculcates the actor; it renders the actor more blameworthy. But given that renunciation/abandonment may serve to negate the requisite mens rea,²² contriving to renounce/abandon may have the opposite effect. Moreover, the more the actor contrives to renounce/abandon, the less firm and the less resolute the actor's intention. Thus the more the actor contrives, the less blameworthy the actor. So, is an actor who contrives to avoid liability by contriving to obtain a renunciation/abandonment defense more or less blameworthy by virtue of the contriving?

²² See *supra* notes 19-21 and accompanying text.

* * *

Under the existing scope of the *actio libera* doctrine, contriving to obtain a defense is inculpatory but contriving to avoid committing an offense is exculpatory. Is there a principled explanation why? Three different rationales were considered and rejected. The rationale of wrongfulness/unlawfulness fails because both an actor contriving to obtain a justification defense and an actor contriving to avoid committing an offense equally contrive to engage in permissible and lawful conduct. The harm rationale also fails because each type of contriving actor may commit the same harm. Moreover, if there is any difference in the type of harm committed, the actor contriving to avoid committing an offense may commit the greater harm than the actor contriving to obtain a defense.

The Legality principle may provide an explanation but only at the cost of demonstrating that the doctrine is arbitrary. If the Legality principle limits the scope of the doctrine's application to those defenses that are not contained in offenses, the doctrine itself depends on the offense-defense distinction many find arbitrary. To avoid this arbitrariness, the scope of the doctrine's application could apply to defenses regardless of whether they are contained in offenses. But then the Legality principle fails to explain why contriving to avoid satisfying some offense elements are inculpatory (those pertaining to defenses) and contriving to avoid satisfying other offense elements (those not pertaining to defenses) are exculpatory. And thus the Legality principle fails to explain why the *actio libera* doctrine is

limited to contriving about defenses. As a result, the doctrine's limited scope is arbitrary. Either way, the *actio libera* doctrine is arbitrary.

The arbitrariness could be avoided in either of two ways—expand the doctrine to include offenses or eliminate the doctrine entirely. Expanding the doctrine to include offenses has some appeal and plausibility. For example, the willful ignorance rule of mens rea might be viewed as an attempt to inculcate contriving about the non-satisfaction of offense elements. However, distinguishing between culpable contrivers and honest, innocent law-aiders—each of whom try to avoid committing offenses—would be exceedingly difficult. Because of this difficulty, the other option of eliminating the doctrine entirely is comparatively more plausible. In addition, the permissibility of contriving may draw some unlikely support. Despite tacitly endorsing the *actio libera* doctrine,²³ in general Leo Katz has passionately defended the permissibility of actors and lawyers contriving and strategizing to follow the letter of the law while perhaps circumventing its spirit.²⁴

²³ See Finkelstein & Katz, *supra* note 1 (defending a particular version of the *actio libera* doctrine).

²⁴ LEO KATZ, *ILL-GOTTEN GAINS: EVASION, BLACKMAIL, FRAUD, AND KINDRED PUZZLES OF THE LAW* 131 (1996) (“At heart, I believe the lawyer is the latter-day incarnation of the Jesuit or Talmudist. His specialty is the ability to generate strategies for capitalizing on the deontological properties of legal rules.”).

Is not the contriver of defenses doing nothing more than “avoision”²⁵ of liability which Katz champions as “a perfectly moral thing to do”?²⁶

But given the nearly universal intuitive appeal of the *actio libera* doctrine, eliminating the doctrine entirely is perhaps also untenable. Thus the challenge remains to supply a principled rationale for the limitation on the scope of the doctrine to defenses. This challenge is even more difficult in light of the many defenses that are explicit or implicit negative elements of offenses.

CONCLUSION

²⁵ *Id.* Based on the distinction in tax law between lawful tax avoidance and unlawful tax evasion, Katz coins the term “avoision” to capture the general nature of conduct “hovering in the limbo between” lawful conduct that complies with both the letter and the spirit of the law and unlawful conduct that violates at least the letter of the law. *Id.* at ix-x. Avoision might be understood as complying with the letter of law while circumventing its spirit.

²⁶ *Id.* at 131.