

# BOSTON UNIVERSITY SCHOOL OF LAW

WORKING PAPER SERIES, PUBLIC LAW & LEGAL THEORY  
WORKING PAPER NO. 02-11



**DOES PUNISHMENT FOR “CULPABLE INDIFFERENCE” SIMPLY  
PUNISH FOR “BAD CHARACTER”?:  
EXAMINING THE REQUISITE CONNECTION BETWEEN MENS REA  
AND ACTUS REUS**

KENNETH W. SIMONS

This paper can be downloaded without charge at:

The Boston University School of Law Working Paper Series Index:

<http://www.bu.edu/law/faculty/papers>

The Social Science Research Network Electronic Paper Collection:

[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=xxxxxx](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=xxxxxx)

Does punishment for “culpable indifference” simply punish for “bad character”?  
Examining the requisite connection between mens rea and actus reus

Kenneth W. Simons\*  
Draft: August 22, 2002

Introduction

- I. The problem, in more detail
- II. Connection requirements for conventional mental states
  - A. Connecting the actor’s purpose or desire with his acts
  - B. Connecting the actor’s knowledge or reckless belief with his acts
  - C. Connecting the actor’s negligence with his acts
- III. Connection requirements for culpable indifference
  - A. Introduction: two culpable indifference criteria
  - B. The cognitive counterfactual criterion: further analysis
  - C. The idealized criterion: further analysis
    1. A single idealized criterion of culpable indifference
    2. More specific multiple criteria of culpable indifference
- IV. Conclusion

Introduction

Is an actor’s “culpable indifference” to the interests of others a proper basis for criminal punishment? Well, it depends.

A. The September 11 Celebrants

Following the September 11 terrorist attacks on New York and Washington, D.C., the FBI reportedly intercepted telephone calls in which individuals celebrated the attacks and exulted in the resulting death and destruction.<sup>1</sup> Such gloating is of course morally reprehensible, but, without more, it is hardly the proper basis for criminal punishment.

B. The Remorseless Negligent Driver

Suppose that a distracted driver, engrossed in conversation with a passenger, accidentally strikes and kills a pedestrian. The driver reacts to the tragedy with utter indifference. His chilling reaction is grounds for moral reproach, but it does not elevate the seriousness of the crime (even if his initial conduct is criminal).

But contrast the following four cases.

---

\* Professor of Law and M. L. Sykes Scholar, Boston University School of Law. © 2002. All rights reserved. I thank Larry Alexander, Kim Ferzan, Gerry Leonard, and Alan Michaels for their helpful comments, and Jason Hall for his valuable comments and research assistance.

In this article, I have willfully ignored Professor Ferzan’s contribution to this symposium. By mutual agreement, we concluded that commenting on each other’s contributions to this forum would be too problematic. I do expect, however, to respond to her detailed critique on another occasion.

<sup>1</sup> “A Nation Challenged: The Suspects; Jubilant Calls on Sept. 11 Led to F.B.I. Arrests,” New York Times, page 1, column 2, October 28, 2001.

### C. The Torturer

Suppose a torturer seriously injures and then abandons a victim, not caring whether the victim lives or dies.<sup>2</sup> The victim does die. Such an actor is highly culpable, and arguably deserves a punishment as serious as one who intentionally kills. Indeed, he deserves a more serious punishment than an intentional killer who acts out of a benign (but not legally justified) desire to relieve the victim's suffering.

### D. The Russian Roulette Player

For the sheer fun of it, a person plays a game of Russian Roulette with an involuntary victim. (Suppose he places one bullet in a gun with six chambers, and credibly commits to pulling the trigger only once.) Such an actor is highly culpable, and arguably deserves a punishment as serious as one who intentionally kills, even though he only poses a one in six chance of causing the death. Moreover, he deserves a much more serious punishment than a person who creates a one in six chance of death for a socially acceptable reason, under circumstances in which the risk-creation is nevertheless unjustifiable. (Suppose a doctor performs a risky procedure on a patient that the doctor believes is necessary to avert the patient's imminent death, but the doctor is grossly negligent in failing to realize that an alternative, much safer procedure would be just as effective.)

### E. The Indifferent Speeding Driver

A speeding driver endangers numerous pedestrians and reveals that he does not care whether or not he causes injuries. (Suppose the driver runs several lights. At each, he is aware that he has barely missed hitting a pedestrian. He never applies his brakes. His passenger urges him to slow down, but he maintains his speed at the final light, where his car strikes and injures a pedestrian.) Such a driver is highly culpable, and arguably deserves more serious punishment than a speeding driver (call her "E2") who creates a similar risk of injury only after attempting to avoid hitting a pedestrian.

---

<sup>2</sup> See *Tison v. Arizona*, 481 U.S. 137, 157 (1987):

A narrow focus on the question of whether or not a given defendant "intended to kill" ... is a highly unsatisfactory means of definitively distinguishing the most culpable and dangerous of murderers. ... [S]ome nonintentional murderers may be among the most dangerous and inhumane of all -- the person who tortures another not caring whether the victim lives or dies, or the robber who shoots someone in the course of the robbery, utterly indifferent to the fact that the desire to rob may have the unintended consequence of killing the victim as well as taking the victim's property. This reckless indifference to the value of human life may be every bit as shocking to the moral sense as an "intent to kill."

However, the Court's glib assertion that the robber's "utter indifference" is as culpable as an intent to kill depends on the nature of the robber's acts and on his state of mind. The culpability should not be deemed equivalent, for example, if the robber believed the gun was unloaded and it accidentally discharged.

#### F. The Indifferent or Wilfully Blind Statutory Rapist

An actor has sexual intercourse with a young person under the age of consent and reveals that he does not care whether or not his victim is under the age of consent. (Suppose he credibly states that he gave the issue no thought whatsoever. Or suppose he realizes that she might be under age but declines to ask her how old she is.) This actor is arguably more culpable than a second actor (F2) who thoroughly inquires about the victim's age but believes a highly implausible story suggesting that the victim is not under age. The first actor might be more culpable even if the actor should, in each case, be aware of an equally significant risk that the victim is under age.<sup>3</sup>

Can these distinctions be defended? In determining what punishment, if any, is deserved, can the law legitimately consider the "uncaring" or "culpably indifferent" attitude of the torturer, the Russian Roulette player, the indifferent driver, or the indifferent statutory rapist, even though it cannot legitimately consider the callous, pitiless, or malevolent attitude of the September 11 celebrants or the remorseless driver?<sup>4</sup>

Antony Duff, Jeremy Horder, Alan Michaels, Samuel Pillsbury, this author, and others have endorsed the criminal law's use of some version of culpable indifference,<sup>5</sup> in addition to the more conventional mental states of

---

<sup>3</sup> In many jurisdictions, no culpability is required for statutory rape. But in a jurisdiction that does require culpability, and in particular a culpability greater than negligence, the distinction between the two actors could matter.

<sup>4</sup> Actually, the criminal justice system does give some weight (for example, in sentencing) to whether the defendant expresses a remorseful rather than defiant attitude towards the crime, apart from whether his attitude influenced his action. But the weight is quite modest. I do not carefully explore the issue in this paper. See note 18= infra.

<sup>5</sup> See R.A. Duff, Intentions, Actions, and Criminal Liability (1990); Jeremy Horder, Gross Negligence and Criminal Culpability, 47 U. Toronto L.J. 475, 508 (1997); Alan Michaels, Acceptance: The Missing Mental State, 71 S. Cal. L. Rev. 953 (1998); Kenneth W. Simons, Rethinking Mental States, 72 B.U. L. Rev. 463 (1992); Kenneth W. Simons, Culpability and Retributive Theory: The Problem of Criminal Negligence, 5 J. Contemp. Legal Issues 365 (1994); Samuel Pillsbury, Judging Evil: Rethinking the Law of Murder and Manslaughter 161-188 (1998). See also Guyora Binder, Meaning and Motive in the Law of Homicide, 3 Buff. Crim. L. Rev. 755, 766-774 (2000) (Review essay); George Fletcher, Rethinking Criminal Law 445-446 (1978) (discussing the German doctrine of "dolus eventualis," which resembles culpable indifference (see Simons, Rethinking Mental States, id. at 489-490)).

The English courts also endorsed a noncognitive conception of recklessness in Caldwell, [1982] AC 341, and Lawrence, [1982] AC 510. In Caldwell, the House of Lords defined recklessness (in part) as doing an act that creates an obvious risk of harm, when the actor "has not given any thought to the possibility of there being any such risk." English commentators are divided on the wisdom of this conception. See Glanville Williams, Textbook of Criminal Law 96-114 (2<sup>nd</sup> ed. 1983) (critiquing the English cases); Duff, id., at 142-179 (offering qualified support). In recent years, the Caldwell rule has been significantly restricted, though it still applies to a few crimes, and a variation of the rule (namely, the "couldn't care less" standard) applies to nonconsent in rape. See Andrew Ashworth, Principles of Criminal Law 183-191, 200-201 (3<sup>rd</sup> ed. 1999).

The Israeli criminal code defines mens rea as to result as either intention or recklessness, and defines the latter, not in terms of awareness of risk, but in terms of indifference:

20. (a) *Mens rea* means awareness ... of the possibility of the consequences of the act being brought about, ... and as regards the consequences, means also one of the following:

(1) "intention" to bring about the consequences;

purpose, knowledge, recklessness, and negligence.<sup>6</sup> Culpable indifference sometimes describes a modestly culpable mental state, sufficient for manslaughter liability (or, with respect to a circumstance element, roughly equivalent in seriousness to cognitive recklessness).<sup>7</sup> It sometimes also identifies a more aggravated form of culpability, sufficient for murder (or, with respect to a circumstance element of an offense, roughly equivalent in seriousness to knowledge).<sup>8</sup> In this article, I will not fully rehearse the ongoing debate about whether, in principle, culpable indifference is indeed a distinct form of fault or culpability, and whether, in practice, it can be defined with sufficient precision to allay concerns about fair notice to potential defendants. Rather, I focus on one important objection to employing culpable indifference as a basic culpability term in defining and grading crimes. The objection is that punishing those who display culpable indifference punishes for “character” rather than for “acts.”<sup>9</sup>

Punishment for culpable indifference, critics maintain, is no more justifiable than punishment for free-floating mental states (for example, punishing a pure bystander who delights in another’s commission of a crime). By contrast, punishment for acts accompanied by the more conventional mental states of purpose, knowledge, and recklessness supposedly does not implicate the “punishment for character” problem. (For convenience, I will identify these critics as “subjectivists.”<sup>10</sup>)

---

(2) "Recklessness," being one of:

- (i) "indifference" to the possibility of bringing about the consequences;
- (ii) "rashness" - assumption of an unreasonable risk as to the possibility of bringing about the consequences hoping that it will be possible to prevent them.

Penal Law of Israel (626/1996) (Unofficial English Translation, "Penal Law--Draft Proposal and New Code," 30 Israel Law Review 5-27 (1996), reported at <http://wings.buffalo.edu/law/bclc/israeli.htm>).

<sup>6</sup> The last of these, “negligence,” is more aptly described as a culpability term, not as a “mental state.”

<sup>7</sup> See Duff, *supra* note 5=; Simons, *Culpability and Retributive Theory*, *supra* note 5=; Pillsbury, *supra* note 5=, at 184. An example of culpable indifference as to a circumstance is the Indifferent Statutory Rapist, Example E, above.

<sup>8</sup> See Michaels, *supra* note 5=, at 1002-1019; Pillsbury, *supra* note 5=, at 184.

<sup>9</sup> Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 Cal. L. Rev. 931, 937-938 (2000); Kimberly Kessler Ferzan, *Opaque Recklessness*, 91 J. Crim. L. & Criminology 597, 620-627 (2001). See also James Brady, *Recklessness*, 15 L. & Phil. 183, 196-198 (1996); Heidi Hurd, *Why Liberals Should Hate “Hate Crime Legislation,”* 20 Law & Philos. 215, 222-226, 229-232 (2001); J.C. Smith & B. Hogan, *Criminal Law* 71(8<sup>th</sup> ed. 1996). In *People v. Phillips*, 414 P. 2d 353, 363-364 (Cal. 1966), Justice Tobriner expressed concern that a jury instruction employing the language “abandoned and malignant heart” from the murder statute could lead the jury to equate the malignant heart with an evil disposition or a despicable character; the jury, then, in a close case, may convict because it believes the defendant a “bad man.” We should not turn the focus of the jury’s task from close analysis of the facts to loose evaluation of defendant’s character [footnote omitted].

<sup>10</sup> “Subjectivists” are those who believe that criminal law mental state categories should be limited to either cognitive states, or mental states of purpose or intention. For discussion, see Ashworth, *supra* note 5=, at 87-88, 160-163; Duff, *supra* note 5=, *passim*.

Subjectivists oppose criminal punishment not only for culpable indifference, but also for negligence. See, e.g., Alexander, *supra* note 9=, at 949-952; Ferzan, *supra* note 9=, at 622. However, the “punishment for character” objection is not the usual reason for objecting to negligence liability; rather, the objection is typically that the merely negligent actor is not sufficiently culpable. And those who endorse the criminal law’s use of culpable indifference typically consider it to be a more serious form of culpability than negligence. See Duff, *supra* note 5=; Simons, *Culpability and Retributive*

The paradigm cases of criminal responsibility, on this view, are where a person acts for the purpose of bringing about the relevant harm, or where he acts notwithstanding his knowledge that the act has a significant chance of causing the harm. In both types of cases, it seems, the actor's mental state is sufficiently connected to the act that the "punishment for character" worry does not arise, because the "connection" requirement in such cases is pretty straightforward. Apart from very unusual problems of lack of "concurrence," it is normally obvious how such mental states justifiably "connect" to the relevant criminal act, and such a connection normally avoids the peril of punishing an actor for mere attitudes or dispositions.

Thus, to be liable for "purposely" causing the death of another, one must simply take some action (such as pulling the trigger of a gun) with the conscious object of causing death. To be liable for "knowingly" causing a death, one must simply take an action with the accompanying belief that death is very likely to result. To be liable for "knowingly" transporting cocaine, one must engage in the act of transporting a package with the accompanying belief that the package contains cocaine. And "recklessness" is analyzed similarly, except that one need only believe that there is a "substantial" rather than "very likely" probability that the harm will occur, or that the circumstance exists (e.g. the package contains cocaine).<sup>11</sup>

But the mental state of culpable "indifference" seems to create a special problem in this regard. To be culpably indifferent, it seems, is either not to care, or to care less than one should, about the legally recognized interests of others in their life, limb, autonomy, or property. Yet it is not at all obvious how such a mental state must "connect" with one's actions. In terms of the introductory examples, why do the first two examples fail to satisfy a connection requirement, and why do the next four examples satisfy it?

This paper attempts to answer that question. And, to see how challenging the question can be, consider one more introductory example, from Larry Alexander, offered as a critique of my own claim that in some cases (such as the last four examples), punishment based on the culpable indifference of the defendant is justifiable.

#### G. Deborah

Suppose ... that Deborah is engaging in her Sunday driving within the speed limit and believes she is creating a low level of risk to others. And suppose that the pleasures of Sunday driving are sufficient to justify imposing that low level of risk on others. But suppose also that Deborah is callous towards the risks she imposes on others, and that even if she were to come to believe that the risks imposed by her Sunday driving were much higher than she actually believes them to be, she would continue to drive notwithstanding that belief. On my account and

---

Theory, *supra* note 5=. For both of these reasons, this essay contains little discussion of negligence as a basis of criminal liability.

<sup>11</sup> Here, and throughout the paper, I follow the Model Penal Code's largely cognitive criterion of recklessness. MPC 210.2(1)(b) (requiring, *inter alia*, that actor subjectively believe that she is creating a substantial risk). If one is employing a different criterion, such as (a) awareness of any risk (even an insubstantial risk) of harm, or (b) lack of belief that the risk is insubstantial, then the analysis throughout can be modified accordingly.

the law's, Deborah is not acting recklessly at this moment and is therefore not culpable. She has an undesirable character trait, one that might dispose her to act recklessly if the circumstances were right, but her counterfactual culpability does not translate into present culpability.

Does Simons believe that because of her callousness, Deborah is now acting culpably, even though the risk level she perceives is justified by her reason for acting? Expanding criminal liability to include callous Deborah would convert criminal liability from being act-focused to being character-focused. Deborah has not acted wrongly; rather, she has a character trait that at most makes her a counterfactual wrongdoer.<sup>12</sup>

Alexander's conclusion is correct: the criminal law cannot justifiably punish Deborah. (The example will be analyzed further below.<sup>13</sup>) But it does not follow that the criminal law can never justifiably employ culpable indifference as a criterion of fault. This paper explores when culpable indifference is indeed especially problematic in punishing merely for an attitude disconnected from conduct, and when it is not. The connection problem, we will see, is much more manageable on some formulations of culpable indifference. At the same time, this problem is hardly unique to culpable indifference; connection problems arise to a surprising extent with the conventional mental states of purpose, knowledge, and recklessness, as well.

A word or two about scope: the analysis in the paper presupposes that criminal punishment—or at least its categorization of culpability—is justified by a retributive rationale.<sup>14</sup> However, the idea of culpability is not restricted to the blameworthiness displayed by the deliberate, conscious choices of the actor; for purposes of this paper, it is an open question whether, for example, culpability for negligence, or for preconscious awareness of risks,<sup>15</sup> or for failure to infer the significance of risk from the facts of which one is aware, could suffice for retributive blame.<sup>16</sup> The analysis also assumes that culpable indifference, like more conventional mental state categories, is defined with respect to a particular element of a crime—for example, the question is whether the actor was culpably indifferent towards causing a death, or towards the victim's nonconsent in rape, or the victim's age in statutory rape (and not whether the actor displayed "culpable indifference" in a more diffuse sense<sup>17</sup>). Further, the discussion focuses on the relevance of culpable indifference as an element of a crime, not as

---

<sup>12</sup> Alexander, *supra* note 9=, at 938.

<sup>13</sup> TAN 143-150= *infra*.

<sup>14</sup> Some of the analysis might apply to deterrent or utilitarian rationales, but I do not explore the possibility here. Cf. Simons, *Rethinking Mental States*, *supra* note 5=, at 503-515.

<sup>15</sup> See Ferzan, *supra* note 9=, at 627-645.

<sup>16</sup> See generally Simons, *Culpability and Retributive Theory*, *supra* note 5=, at 383-384.

<sup>17</sup> For exploration of the latter conception, see Simons, *Rethinking Mental States*, *supra* note 5=, at 488-489 n. 89 (distinguishing "local" or element analysis of mental states from a more diffuse, "global" analysis); Simons, *Culpability and Retributive Theory*, *supra* note 5=, at 396. The language of the Model Penal Code's extreme indifference murder provision is ambiguous in this regard: "extreme indifference to the value of human life" (rather than, say, "to the prospect of causing the victim's death") could be interpreted as expressing the global rather than local conception.

a possible mitigating or aggravating factor in sentencing.<sup>18</sup> Finally, the analysis addresses completed crimes, not attempts.<sup>19</sup>

## I. The problem, in more detail

The mental state of the actor is crucially important in determining both whether the actor is criminally responsible, and how serious a crime he has committed. A number of different mental states are generally viewed as relevant to both questions. The reigning hierarchy in contemporary American law is that proposed by the influential Model Penal Code, which distinguishes (in increasing order of culpability) negligence, recklessness, knowledge, and purpose.<sup>20</sup> Other possibly relevant mental states include motives (whether exculpatory or inculpatory), willfulness,<sup>21</sup> dishonesty,<sup>22</sup> wilful blindness, hope, different varieties of negligence,<sup>23</sup> and, of course, (simple) culpable indifference and (extreme) culpable indifference or “depraved heart.” Moreover, such emotions as anger, fear, and remorse also play a role in criminal grading, though less directly.<sup>24</sup>

---

<sup>18</sup> There might well be good reasons (grounded in retributive or other principles) to relax “connection requirements” at sentencing and to admit, as an aggravating factor, evidence of culpable but “disconnected” attitudes, desires, or emotions (such as gloating about a crime); or to admit, as a mitigating factor, evidence of some “disconnected” attitudes such as remorse.

Of course, gloating or lack of remorse expressed only after a crime has been committed can have evidentiary significance: for example, it makes more probable the inference that at the time of the crime itself, the actor culpably failed to take advantage of opportunities to avoid or mitigate the harm, or that he intended (or at least recklessly caused) the harm. And emotional reactions occurring after the actor has committed a crime or attempt can have independent substantive significance in very limited circumstances. Perhaps the actor had an opportunity to alleviate or mitigate the effects of the harm he risked or caused. Or perhaps the victims become aware of the actor’s response, and this mitigates or aggravates the emotional harm to victims, just as brutally causing harm to V1 in the presence of V2 justifiably aggravates the seriousness of the harm.

To be sure, on a consequentialist account, gloating or lack of remorse might demonstrate future dangerousness. But this paper assumes a retributivist approach.

<sup>19</sup> Inchoate crimes raise special difficulties. The requisite connection between the actor’s mental state and her preparatory or dangerous acts (and between her mental state and the ultimate crime) depends on the special policies that supposedly justify inchoate criminal liability. I don’t address these issues here, except briefly. See TAN 148-149= supra.

<sup>20</sup> Model Penal Code §2.02 (2) (1985).

<sup>21</sup> Although willfulness sometimes means purpose or knowledge with respect to a fact made relevant by the offense, it sometimes has a distinct meaning, such as awareness that one is violating the law. See Joshua Dressler, Understanding Criminal Law 127-128 (3<sup>rd</sup> ed. 2001).

<sup>22</sup> See Don Stuart, Canadian Criminal Law: A Treatise 234-236 (4<sup>th</sup> ed. 2001); *United States v. Farrell*, 126 F.3d 484 (3<sup>rd</sup> Cir. 1997)(analyzing federal witness tampering statute requiring government to prove an intent to “corruptly persuade”).

<sup>23</sup> A proposed codification of British criminal law differentiates three types of negligence—“heedlessness” (akin to Model Penal Code negligence), gross negligence, and simple carelessness. Law Commission No. 143, Codification of the Criminal Law: A Report to the Law Commission 183-184 (Proposed Draft, 1985). And the conceptions of negligence recognized in American jurisdictions also reflect some diversity. See Wayne LaFare, Criminal Law 249-252 (3<sup>rd</sup> ed. 2000).

<sup>24</sup> Anger and fear are explicitly relevant to provocation. (If somewhat excusable, they can result in mitigation of a murder charge.) Fear is implicitly relevant to duress and self-defense. (Unless the actor asserting duress or self-defense is acting out of fear, he is quite unlikely to obtain the benefit of

But an actor's mere possession of one of these legally relevant mental states hardly suffices for criminal liability. Free-floating desires, intentions, beliefs, or attitudes, without more, do not justify criminal liability. In addition, the mental state must be "connected" to the relevant criminal act or omission in the right way. (I employ the vague term "connection" deliberately, for at this point, I do not want to presuppose what kind of relationship a mental state must have to the criminal act.)

This "connection" requirement is expressed in a number of different criminal law doctrines, which can be placed within two broad categories. The first category of such doctrines involves inculpatory mental states. And here, a range of doctrines expresses this basic idea: neither the harboring of a "bad" intention nor the possession of any other inculpatory mental states suffices for criminal liability. Thus, the act requirement forbids punishment for mere thoughts or bare intentions. The voluntariness requirement forbids punishment for bodily movements that another person initiates or controls, or that are not within the potential conscious control of the actor—even if, as it turns out, the "actor" also had an inculpatory mental state at the time of his involuntary conduct. The concurrence requirement provides that the mens rea and actus reus must "concur" in time or in causal interaction (as we shall soon see). Finally, an actor cannot be convicted merely for a status, such as being addicted to drugs.<sup>25</sup>

Consider some examples. Pursuant to the act requirement, D is not liable for any crime even if he forms the firm intention to kill V, if he takes no action to execute that intention. Similarly, D is not liable for a crime merely because he believes that a particular package he happens to observe contains illegal drugs, unless he acted to purchase, transfer, or obtain possession of those drugs.<sup>26</sup> Pursuant to the voluntariness requirement, D is not liable for murder if E hypnotizes D into killing V,<sup>27</sup> or if E uses D's body as an instrument to kill V<sup>28</sup>—and this is so even if D is glad to see V dead.

Under the concurrence requirement, the culpable state of mind must "concur" with the act causing the harm in time, and in the right way. (I will say more about this requirement below.) Thus, Bill would not be liable for intentional murder in the following example:

Suppose [Bill] is out driving thinking about how he is going to kill his uncle, and suppose his intention to kill his uncle makes him so nervous and excited that he accidentally runs over and kills a pedestrian who happens to be his uncle.<sup>29</sup>

---

either defense.) Remorse is implicitly relevant to renunciation of attempts. (If the actor's renunciation is not motivated by remorse, he might have difficulty establishing this defense.)

<sup>25</sup> The prohibition on status crimes is, to some extent, constitutionally grounded. See Alan Michaels, *Constitutional Innocence*, 112 Harv. L. Rev. 828, 877 n. 270 (1999).

<sup>26</sup> However, even if he did not knowingly receive the package that contained illegal drugs, his mere omission to rid himself of the package can result in criminal liability for possession, if he "was aware of his control thereof for a sufficient period to have been able to terminate his possession." Model Penal Code §2.01(4).

<sup>27</sup> This is the Model Penal Code's unqualified position, though some jurisdictions might take a more qualified view. See Dressler, *supra* note 21=, at 87, 97; LaFave, *supra* note 23=, at 209.

<sup>28</sup> Suppose E pushes D against V, causing D's body to knock V off a ledge to V's death.

<sup>29</sup> John R. Searle, *Intentionality: An Essay in the Philosophy of Mind* 82 (1983). I examine a similar example ("Jane") below. See Ferzan, *supra* note 9=, at 622-623, discussed at notes 135-140= *infra*.

Or, if D now knows that the package he delivered for E contained illegal drugs, but came to this realization only after delivering the package, he is not guilty of knowingly transporting illegal drugs.

Second, the “connection” requirement is also expressed in a converse set of doctrines: “good” intentions and other exculpatory mental states do not suffice to provide a criminal law defense or partial defense. Again, the mental states must be appropriately expressed or effectuated in action.<sup>30</sup> Thus, renunciation or abandonment of a crime is not a defense, except in very limited circumstances, and except for the inchoate crimes of attempt, solicitation, and conspiracy.<sup>31</sup> Accordingly, D remains liable if, after committing a (non-inchoate) crime, he regrets his action. In the case of larceny, even return of the property to the owner does not preclude criminal liability.<sup>32</sup> Moreover, with respect to defenses and negative elements of offenses, an arguably exculpatory mental state might not suffice. A rapist who is aware that the victim probably does not consent is guilty even if he hopes that she does consent. An actor who maliciously shoots V as V walks into a room is not justified by self-defense even if, unknown to the actor, V was about to shoot him, and even if, had the actor known this, he would have shot V in justifiable self-defense.<sup>33</sup> And an actor who feels the emotional effects of provocation (such as fear or anger) only after killing the victim, or only so much earlier than the killing that his passion has had a chance to “cool,” might not be entitled to have a murder charge mitigated to manslaughter.<sup>34</sup> Thus, a concurrence requirement applies to material elements of defenses as well as offenses.

What justifies these connection requirements? Several explanations exist. For example, agents whose culpable mental states are “disconnected” from their actions are sometimes, for that reason, not very dangerous. And extending

---

<sup>30</sup> An important qualification here: on some views, one is entitled to a defense even if one lacks an exculpatory mental state. On this approach, the unknowingly justified actor is entitled to a defense, and at worst is liable for an attempt if, under the facts as he believed them to be, he (a) is committing a crime and (b) lacks a defense. However, on any approach under which an exculpatory mental state is required, an appropriate connection of that mental state to the relevant defense is also required.

<sup>31</sup> See, e.g., Model Penal Code §§5.01 (4)(attempt), 5.02 (3)(solicitation), 5.03 (6)(conspiracy); Dressler, *supra* note 21=, at 404-406 (noting that some common law courts recognize renunciation defense for attempt). Moreover, even if it otherwise recognizes a renunciation defense, a jurisdiction might preclude it when a relevant social harm occurs (such as contribution to another’s commission of a crime, or, in attempt, causation of physical harm to the victim). See, e.g., *State v. Smith*, 409 N.E.2d 1199 (Ind. Ct. App. 1980) (no defense when actor renounces and saves life of intended murder victim after stabbing him twice in the chest). See Dressler, *id.* at 405; Daniel G. Moriarty, *Extending the Defense of Renunciation*, 62 *Temple L. Rev.* 1, 50-54 (1989).

Renunciation doctrines typically require the actor to prevent the crime or otherwise to prevent his own conduct from facilitating a crime. For example, under the Model Penal Code, accomplice liability is precluded if the actor terminates his complicity before the crime is committed, and if he either wholly deprives his help of any effectiveness or makes proper efforts to prevent the crime. Model Penal Code §2.06(6)(c). Thus, an actor who satisfies such a requirement is not merely expressing a form of remorse “disconnected” from the social harm that he has caused, but has actually acted to minimize such harm.

<sup>32</sup> See LaFave, *supra* note 27=, at 812.

<sup>33</sup> I am assuming that justification does require an exculpatory mental state. See note 30= *supra*.

<sup>34</sup> See Dressler, *supra* note 21=, at 532-533. Of course, the “no cooling time” requirement has been softened or abandoned in some modern formulations of voluntary manslaughter.

criminal liability to such cases would broaden police and prosecutorial power to interfere with citizens' liberty. But one very important rationale for requiring an appropriate connection between mental states and acts is to avoid punishing individuals simply because they display a "bad character."<sup>35</sup> On this view, the harsh sanctions of the criminal law should not be brought to bear on individuals who have not yet done anything wrong, but who merely have disreputable—or even dangerous—character traits. Criminal punishment for having a callous attitude towards the suffering of others, or even for delighting in others' suffering, is inconsistent with the autonomy that free citizens deserve in a liberal state. We are similarly, and properly, reluctant to impose punishment on a person simply for being short-tempered, or paranoid about threats, or holding racist attitudes, unless and until these characteristics are expressed in action.<sup>36</sup>

## II. Connection requirements for conventional mental states

The concern about punishing simply for character is therefore legitimate and important. But, before we consider whether the connection requirements for culpable indifference are especially problematic in this regard, it is highly instructive to review the connection requirements for the more conventional mental states—of purpose or desire, of knowledge or reckless belief, and of negligence. These requirements are not as straightforward as they first appear. At the same time, the analysis of these conventional mental states will pay useful dividends when we consider connection requirements for different conceptions of culpable indifference in a later portion of the paper.

### A. Connecting the actor's purpose or desire with his acts

Examination of the required connection between an actor's purpose, intention, or desire and his acts will suggest the following general points. First, the motivational connection requirement for purpose is not especially troublesome. Second, the actor's culpable "desire" (short of intention or purpose) normally fails any defensible connection requirement; accordingly, the Model Penal Code's suggestion that the mental state of "hope" suffices for criminal punishment is objectionable. Third, purpose is often defined counterfactually, especially to clarify whether the defendant acted merely with

---

<sup>35</sup> See sources cited in note 9, *supra*. And correspondingly, we might justifiably be reluctant to reward a person (by way of mitigation or a full criminal law defense) for "good character" if the relevant character trait is not appropriately connected to her acts.

<sup>36</sup> A related concern is that punishing simply for "bad character" in this sense commits the state to an indefensible perfectionist political theory, for the state is attempting, by such punishment, to promote the virtue of its citizens based on a necessarily controversial view of the good. This concern, however, is quite distinct from the concern for autonomy: one might be very troubled by threats to autonomy yet support certain types of perfectionist legislation. A criminal law that avoids punishment for "character alone" and carefully restricts its focus to acts motivated by culpable states of mind could nevertheless express a particular view of the good—or, more precisely, of the "bad." Even if that view has perfectionist elements—for example, insisting that men improve their understanding of what sexual autonomy means to women, that negligent actors take more care and use their capacities to conform to a more ideal standard of behavior, or that murderers especially avoid acting for certain highly immoral motives (or else suffer a greater penalty)—the focus on acts at least avoids the problem of punishing for attitudes alone.

knowledge or instead with purpose. (As we will see, this feature of the conventional mental state of purpose undercuts the complaint that if the more unconventional “culpable indifference” is defined by a counterfactual criterion, then this will permit punishment for character.) Fourth, in carefully defined contexts (including bias crimes, provocation, and negligence), some nonmotivational desires do have a defensible connection to action.

To begin, consider “purpose,” in the narrow sense of either the conscious object of one’s actions (the MPC definition), or the intention with which one acts.<sup>37</sup> How must purpose or intention connect to the relevant action? The analysis is usually straightforward, at least in principle. Thus, if the state wishes to show that D’s acts of aiming a gun at V and pulling the trigger were for the purpose or with the intention of killing V, the state must show:

- (a) When he acted, D believed that his actions could<sup>38</sup> bring about V’s death,
- (b) V’s death is what D desired or planned, and
- (c) D took those actions in order to bring about, or as part of a plan to bring about, V’s death.<sup>39</sup>

The last clause is crucial. Often, “purpose” is casually defined simply as desire,<sup>40</sup> or as desire plus belief.<sup>41</sup> But purpose, properly understood, requires

---

<sup>37</sup> Intention is subtly different from purpose:

*Intention* can mean purpose, but it can also mean a nonpurpose plan of future action. Doris might break into a house with the intention of stealing something inside, but she might not have broken in for that purpose. (Her purpose, instead, was to impress her friends.)

Simons, *Rethinking Mental States*, supra note 5=, at 543 n. 272. But I don’t explore the distinction here, because it is rarely of importance in criminal law.

<sup>38</sup> What is the threshold? If D shoots at V from a mile away, and believes there is only a 1 in 1 million chance he will succeed, can he act “purposely”? Or does he merely “hope” or “wish” to succeed? Perhaps it is enough that the actor believes he will increase the risk of the relevant harm by any amount, however small. See Alexander, supra note 9=, at 942. In any event, it is clear that if D believes that his act has no chance of causing a result, he cannot “purposely” cause the result, and cannot have the “intention” of causing it. If, incredibly enough, his act does cause the result, that result is one he merely desires, wishes, or hopes for. At the same time, however, that might well be enough of a “connection” for legal purposes—but only if the actor took the action in execution of his desire or hope to bring about the consequence. See note 46= infra.

<sup>39</sup> For similar accounts, see Michael Moore, Law & Psychiatry: Rethinking the Relationship 9-14 (1979); A.P. Simester & G.R. Sullivan, Criminal Law: Theory and Doctrine 117 (2000). Whether, and in what sense, “desires” should be part of this type of practical syllogism is a matter of philosophical controversy. For the argument that reasons, rather than desires, should be taken as the more basic notion in practical reasoning, see T.M. Scanlon, What We Owe to Each Other, ch. 1 (1998).

<sup>40</sup> Thus, the Supreme Court has explained that a person intends a result of his act either when he knows the result is practically certain to follow, or “when he consciously desires that result, whatever the likelihood of that result happening from his conduct.” *United States v. United States Gypsum Co.*, 438 U.S. 422, 445, 98 S.Ct. 2864, 57 L.Ed.2d 854 (1978) (quoting Wayne R. LaFare & Austin W. Scott, Jr., Criminal Law 196 (1972)). See also Jeremy Miller, *Mens Rea Quagmire: the Conscience or Consciousness of the Criminal Law?*, 29 W. St. U. L. Rev. 21, 53 (2001).

more than this. Even if, at the very time that he acted, the agent possessed both the relevant belief and the relevant desire (for example, to bring about X), it does not necessarily follow that he took the action for the relevant purpose (i.e., to bring about X). His actions could instead have been taken for a distinct purpose, notwithstanding his possessing such a belief and such a desire. For example, suppose the following:

- (1) Prison inmate D's only purpose in burning the prison guard's office is to escape;
- (2) Due to personal animosity, D also does desire or "hope" that the guard will be injured; and
- (3) D believes that the burning of the office might well cause injury to the guard.

Nevertheless, the following could be the case:

- (4) Injuring the guard is not part of D's purpose.

Thus, suppose that D could have set the fire in such a way as to increase the chance that the guard would be injured, but declined to do so. In such a case, given that D's only purpose in acting was to escape, desire for a result (2) plus belief (3) does not entail purpose or intention.<sup>42</sup> If purpose to cause injury or death to the guard has independent significance in determining D's punishment, then D's malevolent desire to injure the guard is insufficient to show purpose.<sup>43</sup>

---

<sup>41</sup> See, e.g. Paul Robinson, Structure and Function in Criminal Law 43 (1997) ("The essence of the narrow distinction between [purpose and knowledge] is the presence or absence of a positive desire to cause the result.")(emphasis original).

<sup>42</sup> I have stipulated that D's only purpose was to escape. Of course, it is possible for an actor to have dual purposes or intentions in acting. If D acted with both the purpose to escape and the purpose to injure the guard, then his action might be overdetermined: absent the purpose to escape, perhaps he would still have set the fire in order to injure the guard. In my example, however, this is not the case.

<sup>43</sup> The guard example is an elaboration of an example offered by Michael Moore, Placing Blame: A General Theory of the Criminal Law 451 (1997). Of course, D can still be liable for recklessly or (possibly) knowingly causing injury to the guard. But if purposely causing harm is punished more severely, then he has not committed that more serious crime. Similarly, if the minimum mental state of a crime is purpose, then in situations like this last hypothetical, the actor is not guilty.

Consider treason, a crime requiring a purpose to aid the enemy, and not merely knowing assistance. John Walker Lindh, the "American Taliban," allegedly joined the Taliban, and remained with them even after the United States began its military campaign against them. Suppose (although this is quite unlikely) that his only purpose in doing so was the personal goal of obtaining fame and writing a best-seller. Then it could be the case that: (a) he believed his conduct would aid the Taliban, and (b) he desired to aid the Taliban's war effort, but (c) his conduct was not for the purpose of aiding the Taliban. For it might be the case that none of his actions were taken with the purpose to aid their war effort. Rather, all of his actions might have been taken with the goal of maximizing his personal fame. (Suppose he passed up more effective military roles in order to take on roles that would make a better story.)

If this conclusion is unpalatable, it might suggest that, as a matter of policy, the mental state requirement for treason should not be so narrow as "purpose," but should embrace certain forms of knowing tangible assistance. Indeed, Lindh was charged, not with treason, but with several other federal crimes, some of which do prohibit knowing assistance to certain groups. The indictment included charges of: (1) knowingly providing material support to a foreign terrorist organization, or attempting or conspiring to do so, 18 U.S.C.A. § 2339B (West 2002); and (2) willfully violating

Moreover, this argument does not merely show how, on sound interpretive principles, a statutory “purpose” requirement should be defined. For, insofar as that malevolent desire has (by hypothesis) no effect on D’s actions, taking that desire into account in determining D’s punishment is unjustifiable, because it fails any defensible connection requirement. Such a desire has no more relevance than the attitude retrospectively expressed by an actor who, after negligently injuring V, gloats about or takes pleasure in the injury.

Not surprisingly, then, the criminal law rarely permits conviction for a mental state of desire unless that desire is effectuated through the actor’s purpose or intention. One interesting apparent exception, however, is the Model Penal Code’s treatment of “hope.” The Code includes “hope” as one way that a person might act “purposely” towards a circumstance.<sup>44</sup> (The other, far more typical way is “belief.”) Thus, the Code provides that a person can act with “purpose” toward a circumstance without believing (or even recklessly suspecting, or negligently failing to suspect) that the circumstance exists. For example, suppose it is a crime to “purposely [rather than knowingly] receive stolen property.” Then an actor is liable if she buys stolen property and wishes or hopes that it is stolen—even if she is quite certain in her own mind that it is not stolen. And this is so even if her desire or hope that the property was stolen did not affect her actions in any way!<sup>45</sup> For the reasons just outlined, the Code position here, endorsing “hope” as a sufficient mental state for a circumstance element, is very problematic.<sup>46</sup>

---

regulations forbidding supplying services to the Taliban, in violation of 31 C.F.R. §§ 545.204, 545.206(a) and 50 U.S.C. § 1705(b) (West 2002). See U.S. v. Lindh, 2002 WL 1489373, \*2–\*3 (E.D. Va. 2002).

<sup>44</sup> Model Penal Code §2.02(2)(a)(ii): “A person acts purposely with respect to [the attendant circumstances of a crime if] he is aware of the existence of such circumstances or if he believes or hopes that they exist.” British law is apparently similar in this regard. See R.A. Duff, Criminal Attempts 10 (1996). See also note 5= supra (under Israeli law, “rashness” is a reckless state of mind in which the actor hopes that the result does not occur).

<sup>45</sup> Consider another example: the actor hopes that his statutory rape victim is under age, because he obtains a deviant pleasure from intercourse with an under age partner, but he genuinely believes that she is very probably over age. If the statutory rape statute in a Code jurisdiction imposes punishment for “purposely engaging in intercourse with a person under the age of sixteen,” but if his desire that she be under age plays no role in his decision to proceed with intercourse, then it appears that his punishment turns merely on his disreputable attitude towards having illicit sex, an attitude that is entirely disconnected from his actual conduct.

<sup>46</sup> On the other hand, it is possible that this “hope” provision is meant to apply only when both: (1) the actor believes that there is a nonzero chance that the circumstance exists, and (2) that belief was part of the actor’s reason for acting. If this is a correct interpretation, then “hope” is not a disconnected mental state, after all. In the stolen property example, perhaps “hope” is satisfied only when the desire that the property be stolen was part of the reason that the actor received the property. (If he has committed a crime in receiving the property, he will impress his criminal friends.) Cf. Simester & Sullivan, supra note 38=, at 124:

The plain case [where an actor has “intention” with respect to a circumstance element] occurs where D rapes V by intentionally having sexual intercourse with her, *hoping* that she does not consent. In this case D intends “sexual intercourse without consent” in its full sense, because it is part of his purpose not merely to have sexual intercourse with V, but to have it without her consent.

One important feature of intention and purpose is that these mental states implicitly are defined counterfactually. If D shoots a gun in the direction of V for the purpose of killing V, and not merely with the knowledge that V would likely die, then it must be the case that D aims at V's death (so to speak!), either as an end or as a means. If D genuinely has this purpose, then he will normally take actions that execute that purpose. If, however, he consciously neglects an easy opportunity to cause V's death, then (putting aside possible weakness of will) he no longer is acting with the same deadly purpose. Thus, an appropriate criterion for identifying purpose, and differentiating it from knowledge, is whether (in hypothetical and relevantly different circumstances) the actor would continue to take actions to bring about V's death, or instead would fail to take advantage of such opportunities.<sup>47</sup> Would D's immediate or further purposes be satisfied even if V does not die, or do these purposes require V's death? Only in the latter case does D act with the purpose of causing V's death.<sup>48</sup>

---

It is indeed plausible to interpret "hope" (as opposed to a mere "wish") as requiring the first condition above—that the actor believes there is some chance that the circumstance exists. See Searle, *supra* note 29=, at 32. (I thank Larry Alexander for pointing this out.) But I am not at all certain that the statutory language implies the second condition. Accordingly, the connection problem remains.

I have been unable to find any case law on the question whether, and in what context, courts have permitted prosecution for actors who hope (but do not believe) that a statutory circumstance exists.

<sup>47</sup> Antony Duff has proposed a counterfactual test of failure: one who intends a result would regard himself as having failed, at least in part, if the result does not occur. Duff, *supra* note 44=, at 17. See also Susan Uniacke, *Double Effect, Principle of*, Routledge Encyclopedia of Philosophy 120 (E. Craig ed. 1998); Michael Bratman, *Intention, Plans, and Practical Reason* 140-143 (1987) (defining intention, *inter alia*, in terms of a disposition to adjust one's behavior in response to indications of success or failure in bringing about the result); Leo Katz, *Ill-Gotten Gains: Evasion, Blackmail, Fraud, and Kindred Puzzles of the Law* 37 (1996) (we differentiate causing death as a means to a goal from causing it as a by-product "by determining whether the defendant, if he had not thought the death tied to be tied to his goal, would have chosen a different course of action").

It might be possible to offer a noncounterfactual account of intention and purpose. Still, it is notable that many plausible criteria of these mental states are indeed counterfactual—especially those criteria that are invoked to differentiate intention or purpose from knowledge or belief.

<sup>48</sup> Thus, the contract killer needs to succeed in killing in order to obtain his fee; he acts with purpose. However, in the famous trolley problem, where a runaway trolley will kill five on the track ahead unless one turns the trolley onto a spur where it is certain to kill only one worker, an actor who turns the trolley kills the one knowingly but not purposely. (Of course, the fact that the death is merely foreseen, not intended, is not sufficient to render permissible the act of diverting the trolley. See Alison McIntyre, *Doing Away with Double Effect*, 111 *Ethics* 219, 236 (2001).)

To be sure, the counterfactual criterion of purpose or intention has notorious difficulties. One can almost always characterize the actor's actual purposes in such a way that the death of the victim is not strictly required. Even the contract killer doesn't really need to kill the victim; he merely needs to make his employer believe that the victim is dead, so that the killer can get his money. So when the contract killer fires ten bullets into his victim's head, the killer may be indifferent to whether the victim (a) immediately dies, or instead (b) merely appears to die, but miraculously survives and digs himself out of his grave—so long as the victim then takes on a new identity and thus does not jeopardize the killer's fee. For further examples along this line, see Moore, *supra* note 43=, at 454-456; Simester & Sullivan, *supra* note 38=, at 120-121; A.P. Simester, *Moral Certainty and the Boundaries of Intention*, 16 *Oxford J. Leg. Stud.* 445, 456-466 (1996).

One possible implication of such outrageous examples is that "purpose" should not be so narrowly defined. However, the greatly broadened definition that some advocate—expanding purpose to include knowledge—is itself extremely problematic. (For example, it could result in the

Often, of course, the distinction between purpose and knowledge is unimportant to criminal liability. In murder, for example, either mental state suffices in most jurisdictions. But the distinction could differentiate grades of an offense,<sup>49</sup> and in some important contexts, it defines the boundary between criminal and noncriminal conduct. Thus, in most jurisdictions, attempt and accomplice liability require purpose, not just knowledge.<sup>50</sup> And treason, among a few other crimes, also requires purpose. Especially in these cases, where purpose marks the threshold of criminality, we must be able reliably to distinguish from purpose not only cases in which the actor only possesses the mental state of knowledge, but also cases in which the actor merely possesses a culpable desire. In addition to the latter, we should require that the desire be motivational, in order to avoid the problem of punishing for character alone. (For example, attempt liability requires purpose, not merely desire; it would be unthinkable to impose attempt liability simply because an actor desired to cause a crime and happened to have the means at hand to do so, but took no steps towards effectuating that attempt.<sup>51</sup>)

The traditional requirement that mens rea and actus reus must “concur”<sup>52</sup> is almost always applied to the mens rea of purpose or intention. The typical “concurrency” problem arises when the actor has an illicit purpose at some point in time, but not at the time he engages in the acts causing the relevant social harm. “Thus it is not a criminal battery for A accidentally to strike and injure his enemy B, though A, on realizing what has happened, rejoices at B’s discomfiture.”<sup>53</sup> Although the problem is often described as a lack of temporal concurrence, it is really an instance of the broader problem of lack of

---

conviction, as an accomplice, of any merchant who provides the principal with an item that the merchant knows will be used in a crime.) A better response to such examples is to bite the bullet (sorry) and concede that strictly defined purpose should rarely, if ever, be necessary for criminal liability or grading. Rather, either purpose or extreme indifference (or perhaps knowledge, if suitably defined and limited) should suffice—for example, for murder or complicity. But I cannot fully explore the issue in this space.

<sup>49</sup> I say “could” because I have been unable to find an example in which, holding all actus reus elements constant, acting knowingly is a lesser grade of an offense than acting purposefully. See also Douglas Husak, the Sequential Principle of Relative Culpability, 1 *Legal Theory* 493, 504 (noting that the Model Penal Code’s own offense definitions do not fully exploit the Code’s four-part mental state hierarchy in grading degrees of a crime).

<sup>50</sup> More precisely, attempt law requires that the defendant’s preparatory acts be taken with the ultimate purpose of committing a crime, and also usually requires that the defendant act with the purpose of bringing about any result element of the offense. And accomplice liability usually requires that the accomplice act with the purpose of facilitating the principal’s conduct.

<sup>51</sup> It would even be unthinkable if she took steps that would otherwise suffice for the actus reus requirement (e.g., going beyond “mere preparation” to a “substantial step”), and had the desire to commit the crime, but did not take those steps as a way of executing that desire. Thus, suppose Jill brings her new gun, fully loaded, to a bar where she meets Kate. Kate had bet Jill \$20 that she wouldn’t dare bring the gun there and pull it out in public. Jill pulls out the gun, points it across the room, and collects her \$20. At that moment, Jill spies her ex-boss Lloyd across the room and says honestly to Kate, “I really have a desire to kill that bastard.” Jill has not committed attempted murder—even if Lloyd is in direct range of the gun.

<sup>52</sup> See Dressler, *supra* note 21=, at 197-199; LaFave, *supra* note 23=, at 283-292; Paul Robinson, *Criminal Law* 21-219 (1997). In British law, the equivalent doctrine is the “contemporaneity” principle. See Ashworth, *supra* note 5=, at 163-166. In Canada, it is dubbed the “simultaneous” principle. See Stuart, *supra* note 22=, at 359-363.

<sup>53</sup> LaFave, *supra* note 23=, at 284.

motivational connection.<sup>54</sup> After all, in this and many other examples, the actor has the desire or even intention to cause a harm at the very same time that she accidentally and faultlessly causes the harm.<sup>55</sup>

But can a person's desires be sufficiently "connected" to his conduct in some way other than by being executed as purposive or intended action? Indeed they can. However, explaining when the connection suffices in this context is a matter of some intricacy. The remainder of this section addresses the problem.

To start, why not employ a straightforward causal test? One might simply ask: if not for his intention to harm V, or even his mere desire to harm V, would D have acted as he did? But this simple causal test is inadequate. Consider a counterexample, offered by Wayne LaFave, involving intention:

Where [Arthur] intends to kill B and gets drunk for this purpose and then, because of his intoxication, inadvertently runs over B, it could be said that the death was a consequence of the prior intent... But ... more is required than that the death-causing acts would not have occurred but for the prior intent.<sup>56</sup>

Like Bill in the earlier illustration, Arthur does not intentionally kill. Now consider this counterexample involving a culpable desire (rather than an intention): One morning, Clara is suddenly overwhelmed by her hatred of her unloving father. Her emotional state causes her to leave for work much later than usual, which in turn causes her to speed as she drives to work, with the result that she accidentally strikes and kills a pedestrian. Clara clearly does not intentionally kill (nor should her somewhat culpable desire affect her liability for negligent homicide). As these examples suggest, the very broad causal test is inadequate: it permits conviction even where the actor's formulation of a culpable intention, or her failure to control her culpable desire, has only a highly attenuated connection to the harm, or displays an insubstantial degree of culpability.

These counterexamples might suggest that any causal test is too broad unless it is narrowed to require a motivational connection. Yet this conclusion is also too quick. For we can readily identify other connection requirements, intermediate between the extremes of a broad causal test and a narrow motivational requirement, that are defensible, and that avoid the problem of "punishing for character."

Consider three illustrations of an intermediate category. A first illustration is "hate" or bias crimes. Criminal provisions authorizing enhanced punishment for bias crimes can reach at least three different kinds of actors: (1) those who commit a crime at least in part for the purpose of expressing hostility or disrespect for a particular racial group; (2) those who commit a crime out of racial animus or hostility; and (3) those who select victims because of their race. I will focus on the first two subcategories.<sup>57</sup> If members of a white supremacist

---

<sup>54</sup> See LaFave, *supra* note 23=, at 284; see also Dressler, *supra* note 21=, at 197-198 (treating temporal and motivational concurrence as independent requirements).

<sup>55</sup> Recall the similar example of Bill, above. TAN 24= *supra*.

<sup>56</sup> LaFave, *supra* note 23=, at 286 n. 21. (I supply the name "Arthur" for ease of later reference). See also Moore, *supra* note 39=, at 16.

<sup>57</sup> In the third subcategory, where an actor selects one victim rather than another for a crime because of race, the actor might not fall within either of the first two subcategories. Rather, he might act with

group decide to harass blacks in order to further their political goals, they exemplify the first category. If, during a sudden confrontation, a racist individual impulsively strikes a person of a different race and utters a racist epithet while doing so, he might exemplify the second category, not the first. In this last example, expressing a racist world-view might well not have been the reason for the action, just as, in the common case where a person responds angrily to an insult, expressing anger need not be the reason for the heated response. Racist impulses and anger can be causes of action without being their “purpose.”<sup>58</sup>

Only in the first of these three subcategories does the actor’s desire operate as his motive or purpose. Yet a person in the second subcategory also arguably deserves additional punishment, for he has culpably failed to control his racist desires or impulses. (Moreover, many bias crimes probably involve spontaneous and impulsive acts that fall within the second category, rather than the more instrumental acts that fall within the first.)<sup>59</sup>

A second illustration of desires or emotions that are relevant to criminal liability, yet need not be motivational, is the venerable doctrine of voluntary manslaughter. Killings in which the actor was provoked to understandable anger, jealousy, grief, or fear can be mitigated from murder to voluntary manslaughter. Whatever the proper scope of the voluntary manslaughter doctrine, and whichever emotions properly mitigate, the role of these emotions is normally as causal explanations, not purposes, of the actor’s conduct. Thus, suppose the actor killed in anger after witnessing the victim harm his child. Whether mitigation will be allowed does not depend on whether expressing anger or obtaining vengeance was the actor’s purpose in acting. It is enough that an angry, vengeful emotional reaction was triggered by the provocation,<sup>60</sup> and was a proximate causal explanation of the killing.<sup>61</sup>

---

an unconscious racist motive or because of an unconscious racist desire, or he might simply use race as a proxy for another relevant characteristic (such as potential wealth, in the case of theft). Whether bias crimes should reach this third subcategory is especially controversial. See Frederick Lawrence, Punishing Hate: Bias Crimes Under American Law (1999); Lu-in Wang, *The Transforming Power of “Hate”*: Social Cognition Theory and the Harms of Bias-Related Crime, 71 S. Cal. L. Rev. 47 (1997); Lu-in Wang, *The Complexities of “Hate,”* 60 Ohio State Law Journal 799 (1999); Allison M. Danner, *Bias Crimes and Crimes Against Humanity: Culpability in Context*, =this volume.

<sup>58</sup> See Kenneth W. Simons, Book Review: *Social Meaning, Retributivism, and Homicide* (Review article: Samuel H. Pillsbury, Judging Evil: Rethinking the Law of Murder and Manslaughter), 19 Law & Philosophy 407 (2000); Hurd, *supra* note 9=, at 219-222 (arguing that hatred and bias are, not specific intentions, but emotions and dispositional beliefs). Contrast a passive individual who, after much psychotherapy, consciously chooses to react to a stressful provocation by expressing anger. Here, the actor does purposely express his anger, but it is a most unusual case!

<sup>59</sup> This is not to say that criminal liability is defensible whenever a racist desire or attitude causally issues in action. If a driver hears a news story on the radio that expresses sympathy for the economic plight of racial minorities, and the racist driver becomes enraged and distracted, leading to an accident, it is hardly appropriate to convict the actor of a bias crime. The racist desire needs to have a direct causal relation to the social harms that bias crimes distinctively implicate, such as racial insult or devaluing of a racial group.

<sup>60</sup> The requirement that the provocation trigger the reaction is usually understood to include both a subjective and an objective dimension—i.e., the actor must have actually been provoked to anger, in circumstances where a reasonable person would have been provoked.

<sup>61</sup> Hurd, in her argument that punishment for hate or bias crimes amounts to punishment for character, concedes that voluntary manslaughter is a context in which emotional states can exculpate, yet concludes that allowing an emotional state such as bias or hatred to inculcate is problematic. Hurd,

Finally, as a third illustration, consider the relevance of intentions and desires to liability for negligence. Recall Bill, whose intention to kill his uncle caused him to be nervous and excited while driving and thus caused him to accidentally run over the victim (who turned out to be his uncle). Although Bill does not deserve to be punished for murder, he might well be liable for negligent homicide, if a reasonable person undistracted by this culpable intention would not have caused the accident.<sup>62</sup> Now imagine a case (“Edith”) in which a mere desire to kill her uncle has a similar effect: engrossed in conversation with a passenger, Edith fulminates about her hatred of her uncle, and therefore loses concentration on her driving, causing her accidentally to run over the victim. And contrast a case (“Flora”) in which a driver loses concentration because she suddenly hears on the radio that a close friend has died. The law of negligence can justifiably (a) treat Edith’s culpable desire as an inexcusable reason for losing concentration and losing control of the vehicle, but (b) treat Flora’s emotional response as forgivable.<sup>63</sup>

In short, the simple causal test (“Did a culpable desire cause a harm?”) is inadequate, because too far-reaching. But it does not follow that only a motivational test provides a sufficient connection between desires and acts to justify increased (or decreased) criminal punishment. Rather, an intermediate position is defensible: some causal but nonmotivational desires do justifiably bear on criminal liability.

This conclusion—that it is legitimate for criminal sanctions to depend in part on an actor’s desires or emotions even when the role of those desires is not motivational—is quite significant. As we will see, it lends some support to the use of a “culpable indifference” standard for criminal liability.

#### B. Connecting the actor’s knowledge or reckless belief with his acts

Now consider the cognitive mental states of “knowledge” and “recklessness.” Assume that “knowledge” requires belief that a harm (such as a death) is highly likely to result, or that a circumstance (such as the presence of cocaine in a package) is highly likely to exist.<sup>64</sup> And assume that “recklessness”

---

supra note 9=, at 217-218. But she fails to explain why retributive theory can justifiably recognize only a desire or emotional state that exculpates.

<sup>62</sup> The case of Arthur, see note 56= supra, should be analyzed similarly.

British scholar Glanville Williams, a stalwart subjectivist who objects to the use of indifference as a mental state, nevertheless would treat a person who is unaware of a risk only because of anger on a par with a person who is actually aware of the risk. Glanville Williams, *The Unresolved Problem of Recklessness*, 8 *Legal Studies* 74, 86 (1988).

<sup>63</sup> Moreover, even a culpable desire should not affect negligence liability if its effect on the actor’s conduct is too attenuated. Recall Clara, above, whose later decision to speed to work “supersedes” and thus makes legally irrelevant her own prior (and possibly culpable) anger, even though the latter is a cause in fact of her negligent act. TAN 56= supra.

<sup>64</sup> The Model Penal Code defines knowledge as requiring the actor’s (a) belief that it is practically certain that a result will ensue, or (b) awareness of at least a high probability that a circumstance exists. §§2.02(2)(b)ii, (7). Although typically described as a mental state, “knowledge” is more than this; it also contains an actus reus requirement, that the state of affairs the actor believes is likely to occur or exist, actually does occur or exist. (One can’t knowingly bring about a death unless the death occurs; one can’t knowingly possess cocaine unless one actually possesses cocaine.) Thus, strictly speaking, “belief” is the relevant mental state; “knowledge” comprises both this mental state and the actus reus requirement (that the belief is correct).

requires belief that a harm has at least a “substantial” (but not necessarily “highly likely”) chance of resulting, or that a circumstance has at least a substantial chance of existing.<sup>65</sup>

What is the required connection between such cognitive mental states and the relevant action? The analysis that follows will demonstrate several points. First, the basic connection requirement is only a minimal constraint—that the actor possess the relevant belief at the time of the action. Second, the requirement does become problematic in those frequent situations when the actor has only latent awareness of a relevant risk or fact, or when the actor acquires his belief well in advance of the criminal act, and then, because of forgetfulness, distraction, or preoccupation, fails to access that prior belief when he acts. Third, two different counterfactual criteria are plausibly relevant to an actor’s culpability in these situations—a criterion of counterfactual consideration (if the actor had consciously considered the legally relevant question, such as whether he is posing a significant risk of death, what would he have concluded?), and a criterion of counterfactual risk-avoidance (if the actor had been aware of the risk would he still have acted in the same way?). Fourth, the reason why an actor is not fully aware of a risk at the time of action may be relevant; if that reason is culpable, the actor might justifiably be treated, for purposes of punishment, as if he were indeed aware.

Let us turn to the basic connection requirement for cognitive mental states. Interestingly enough, this requirement is quite minimal: ordinarily, the actor need only possess the belief at the time that he acts. So long as D believes, when he shoots the gun in the direction of V, that he is highly likely to kill V, he “knowingly” kills. So long as D believes, when he transports the package containing cocaine, that it is highly likely to contain cocaine, he “knowingly” transports cocaine. Indeed, if we imposed a narrower or more stringent connection requirement—namely, that the actor must have taken the action only because he believed that the action would likely cause death—then the actor would ordinarily have acted for the purpose of bringing about death.<sup>66</sup> On this narrower view, knowledge and recklessness would not be distinct levels of culpability, but would collapse into purpose.<sup>67</sup>

---

<sup>65</sup> The definition of “recklessness” typically also includes a requirement that the risk be unreasonable, or grossly unreasonable, to take. See Model Penal Code §2.02 (2)(c). For now, I focus only on the cognitive or “awareness of risk” component.

“Recklessness” also can refer to culpable indifference, as in the terminology “reckless indifference.” See Simons, *Rethinking Mental States*, supra note 5, at 486-490. In order to keep the concepts distinct, I avoid that terminology in this paper.

<sup>66</sup> “Ordinarily,” but not always. The “because” condition could be causal but not motivational. And then the following is possible: D knowingly causes a death, and would not have caused death but for his belief that death was a very likely result of his actions; yet D did not purposely cause death. For example, D believes that the guard’s death is likely as a result of burning the guard’s office, but D’s only purpose in burning the office is to escape. Yet his belief that the death is likely could still “cause” him to burn the office. Thus, suppose D previously disclosed to E that he was considering this escape plan, but was very unsure about whether to go forward with it. When E heard that the plan was likely to kill the guard, whom E secretly hates, E persuaded D that the escape plan was sensible. E’s secret motivation for persuading D is to get the guard killed; but this need not be D’s motivation. Still, but for D’s belief that the death was likely to result, E would not have intervened, and D would not have caused the death.

<sup>67</sup> The analysis of recklessness is actually a bit more complex than indicated in the text. If the required connection between awareness of a significant (but less than “likely”) risk of death were

The required connection for cognitive mental states, in other words, is implicit. If a person acts despite his belief that he is likely to cause harm, or despite his belief that he has created a significant risk of harm (or if he acts despite his belief that an inculpatory circumstance is likely or is a substantial risk), then his culpability arises from his failure to give sufficient weight to that consideration in his conduct.<sup>68</sup> Of course, if the actor acquires a belief only after he completes the relevant act, the belief fails this “connection” requirement and is irrelevant to criminal liability.<sup>69</sup>

What about beliefs acquired well in advance of the relevant act? This “connection” problem is far more troubling. The difficulty arises when those beliefs are not “operative” at the time of the action. Suppose I walk out of my house with a loaded gun, forgetting that I loaded it last year. Can I be convicted of a crime of knowingly carrying a loaded gun in public? Suppose I loaded it only this morning, but still forgot? On a narrow view of knowledge, I am not liable.<sup>70</sup> On a broader view, once I acquire knowledge, the connection requirement is still satisfied. In effect, I have an ongoing duty not to forget, as well as a duty to employ the knowledge at the operative time.<sup>71</sup>

Moreover, suppose I “know” that the gun is loaded, not in the sense of vivid current conscious awareness of this fact, but in the straightforward sense that I would readily concede the fact if I were asked and were to reply honestly—just as I “know” that 2+2=4, and that George Bush is our current President, even when those beliefs are not conscious, and are not occupying my thoughts. Is it enough, to warrant criminal liability for “knowledge,” that I possess such latent knowledge?<sup>72</sup> Even in a case when the actor ordinarily would be quite conscious

---

narrow in the way just suggested, then proof of recklessness would require proof that the actor took an action only because he believed that the action would create a significant risk of death. But this implies only that the actor intended to create a significant risk of death, not that the actor fully intended to cause death. (Consider a driver who derives a thrill from endangering pedestrians.) Still, the motivational interpretation remains a narrow one, for it would treat as reckless only those actors who intend to endanger others, not the much broader class of actors who are culpably willing to create unjustifiable risks as side effects of seeking other ends (including perfectly licit ones).

<sup>68</sup> “What makes the knowledge morally culpable is her action connected with her knowledge. It is the action in spite of the knowledge.” Michaels, *supra* note 5=, at 967 (emphasis in original).

<sup>69</sup> “The crime of receiving stolen property knowing it to be stolen is not committed when A receives from B property which he does not know to be stolen, though when A later learns it is stolen he decides to keep it.” LaFave, *supra* note 23= at 284.

The proposition in the text is not true, however, if the belief triggers a subsequent duty to prevent an initial act or omission from causing harm. In the prior example, it might be sensible to criminalize an actor’s failure to return stolen property, even if he obtained it innocently.

<sup>70</sup> For discussion, see Eric Colvin, *Recklessness and Criminal Negligence*, 32 U. Toronto L. J. 345, 368-369; Brady, *supra* note 9=, at 188, 199; see also Ashworth, *supra* note 5=, at 185-186 (considering the related question of whether an impulsive, unthinking action can be characterized as “advertent” with respect to a risk that one would recognize if one had given the matter a moment’s conscious thought).

<sup>71</sup> Notice, however, that this broader view presupposes the kind of affirmative duty that critics of criminal negligence liability would reject. See note 10= *supra*. If they wish to maintain their rejection of negligence liability, it seems that they should reject this broader view of knowledge, as well. But that commits them to an extremely narrow view of criminal liability for knowledge or recklessness.

<sup>72</sup> For some discussions, see Alexander, *supra* note 9=, at 953-954 n. 62; Duff, *supra* note 5=, at 164-165; Ferzan, *supra* note 9=, at 628-641; Horder, *supra* note 5=, at 510-513; Simons, *Culpability and Retributive Theory*, *supra* note 5=, at 384-385; John Searle, *supra* note 29=, at 141-159. For an

of a fact—for example, when a bank robber is engaged in a robbery with a loaded gun in his hand, having knowingly loaded the gun earlier the same day—the actor might not bring that knowledge to bear on his action at the relevant time, because of distraction, preoccupation, forgetfulness, or for myriad other reasons.<sup>73</sup>

These complexities reveal that the connection requirement for cognitive states—“acting in the face of an inculpatory belief”—is not nearly as straightforward as first appears. To be sure, the underlying rationale of that requirement is clear enough: criminal liability is appropriate because the actor could and should have decided differently in the face of those beliefs. Moreover, a similar rationale is meant to explain the grading difference between reckless and knowing acts. As the degree of risk that the actor subjectively appreciates increases, his decision to act notwithstanding those risks becomes increasingly culpable. (The knowing actor does proceed with the criminal act despite knowledge, but we usually cannot say whether the merely reckless actor would proceed to act if he had had such knowledge.)

Yet in such cases as forgetting, distraction, and latent knowledge, the actor who at some point in the past possessed a legally relevant belief and now encounters the risk does not necessarily reveal the presumptive culpability characteristic of knowing actors—namely, that their actions show that they give far too little weight to a highly probable harm in their practical reasoning and action. Rather, in such cases, the actor fails (for some reason) to give the risk any weight at all. Is it then proper to treat him as highly culpable? Is he as culpable as an actor who chooses to create a risk with full conscious awareness of that risk?

The answer depends in part on why the actor forgot or was distracted. If the reason is itself culpable, a court would be justified, or at least more justified, in treating the actor as “knowing,” than if the reason is nonculpable. (The wilful blindness doctrine is similar in its rationale, though there, the culpable reason for lack of knowledge essentially converts recklessness into knowledge,<sup>74</sup> while here, a culpable reason for lack of conscious awareness might convert negligence into recklessness or even knowledge.)

For example, suppose it is a crime to knowingly or recklessly sell liquor to a minor. The owner asks a somewhat youthful-looking customer for identification, and she presents a driver’s license that the owner recognizes as forged. In the first variation, the owner is then distracted by a fight occurring on the premises. The owner intervenes and successfully quells the fight after lengthy discussion with the participants. When he returns to the customer, he has forgotten that she presented a forged license, and he sells her the liquor. In the second variation, the owner is (after seeing the apparently forged license) distracted because he is completing the sale of a large amount of Russian vodka whose sale (as he knows) the government has recently banned; again, when he returns to the customer, he sells her liquor, having forgotten about the forged

---

argument expressing doubt that conscious awareness characterizes even cases where the actor has a criminal intention or purpose, see Bruce Ledowitz, *Mr. Carroll’s Mental State Or What Is Meant By Intent*, 38 Am. Crim. L. Rev. 71 (2001). For a response, see Ferzan, *id.* at 648-650. See also Peter Cane, *Fleeting Mental States*, 59 Camb. L. J. 273 (2000); Hyman Gross, *A Theory of Criminal Justice* 90-98 (1979).

<sup>73</sup> See Ashworth, *supra* note 5=, at 187 (on preoccupation).

<sup>74</sup> See generally Michaels, *supra* note 5=, at 977-995.

license. It is defensible to treat the owner as knowing, or at least reckless, in the second scenario but not the first.

Perhaps the “latent knowledge” problem could be solved by employing a rather straightforward counterfactual consideration criterion: if the actor had given the matter a moment’s thought, would she have been aware of the relevant risk? And would she then have characterized the risk as likely, or as significant? If so, then perhaps the actor should be deemed aware of a likely or significant risk of harm, even if she was not consciously attending to such risks at the time of action. But this criterion seems deficient, for it fails to differentiate between culpable and nonculpable reasons for lacking conscious awareness of the sort just noted.

Moreover, an even stronger basis for justifying a moral equivalence between latent and occurrent knowledge employs a counterfactual risk-avoidance criterion. This criterion looks, not just to whether the actor’s reason for not bringing latent knowledge to the forefront was culpable, but also to whether the actor, if consciously aware of the risk, would have avoided it. If he would not have, but would still have taken the risk, then it might be proper to treat him the same as an actor who was actually aware of the risk and took it. Under this criterion, the relevant inquiry is not whether the actor would have been consciously aware of the risk if he had considered the question, but whether, if the actor had brought the knowledge to consciousness, she would have avoided the risk, or instead would have acted the same way. Thus, in the second scenario above of a liquor sale to a minor, an even stronger basis for treating the owner as reckless or knowing is that he would have sold the liquor to the minor even if he had remembered that she presented a forged license. (Suppose that, just prior or just subsequent to his sale to this customer, he did, in fact, sell liquor to a minor with an obviously forged license.)<sup>75</sup>

Of course, there are significant evidentiary difficulties with the application of such a counterfactual test. If this test were employed, the state would need, and would often lack, clear proof of what the actor would otherwise have done. But the central point is that without this or some other supplementary principle, a legal requirement of “belief” will, if applied honestly, be difficult to satisfy in a surprisingly wide range of cases.<sup>76</sup>

The utility of this counterfactual risk-avoidance criterion is not limited to the problem of identifying which cases of latent awareness should be treated as cases of patent awareness of a risk. The criterion also helps explain the common

---

<sup>75</sup> The culpable v. nonculpable reason approach illustrated with the two “sale to minor” scenarios often gives the same results as the counterfactual risk-avoidance approach. If one’s reason for not bringing latent knowledge to the forefront of consciousness is nonculpable, that provides little basis for concluding that, if one had been aware of the risk from one’s behavior, one still would have encountered it. By contrast, if one’s reason for unawareness (notwithstanding latent knowledge) is culpable, then, depending on the particular culpable reason and the facts, it might be clear that the actor would still have encountered the risk if he had noticed it.

<sup>76</sup> Ferzan addresses this problem in part with an interesting proposal to expand the Model Penal Code’s definition of recklessness, so that the definition embraces some actors who are not conscious of the substantiality and unjustifiability of the risk. In essence, she would count as reckless those actors who are conscious of the dangerousness of their conduct, even if they are not conscious of the substantiality of the risk. Ferzan, *supra* note 9, at 644. Whatever the merits and drawbacks of this proposal, it indicates that even orthodox subjectivists are reluctant to require conscious awareness of a risk before imposing criminal liability.

grading distinction between reckless awareness and knowledge, as follows. In the case of both reckless and knowing actors, what is culpable is the actor's failure to take account of the beliefs that she actually possesses, when she acts. That failure can be either moderately or seriously culpable, depending on the seriousness of the risk that she believed she was facing, a risk that she did not appropriately consider or respond to.<sup>77</sup> But why is the reckless actor less culpable than the knowing actor? One explanation is this: we ordinarily cannot be confident that the reckless actor would have created the risk of a deadly harm, or taken the risk of transporting the package, if she knew that the harm was likely to occur or that the package was likely to contain cocaine. The knowing actor who does act in the face of these likely harms or circumstances has ordinarily demonstrated a more serious degree of culpability than the actor who acts only in the face of significant (but less than likely) risks. (And similarly, the inadvertent actor who should have been aware of a risk is ordinarily less culpable than an advertent actor, because we ordinarily cannot be certain that the actor, if he had been reckless or knowing, would still have taken the risk.)

Thus, the grading distinction between different degrees of belief and awareness can be justified by a negative counterfactual criterion: it cannot be assumed that the inadvertent actor would have acted as the advertent (reckless) one did, or that the reckless actor would have acted as the knowing one did. Normally, it is more accurate to assume the contrary, that he would not have so acted.

Should we take the next step, and endorse the positive version of the last test? If a reckless actor would have acted the same way and taken the risk even if he had (like other knowing actors) believed the risk was likely, perhaps he should be punished the same as a knowing actor. And perhaps even an inadvertent actor who would have taken the risk if (like a reckless actor) he had been aware of it should be punished the same as a reckless actor.

This next step is much more controversial, for it premises criminal punishment on the result of a hypothetical inquiry. I postpone full discussion of the issue until later.<sup>78</sup> But, before concluding our discussion of connection requirements for beliefs, it is worth exploring one other context in which a counterfactual criterion is employed in determining culpability for reckless awareness of a risk—namely, the relevance of the actor's intoxication.

In many jurisdictions, if a crime requires reckless awareness of risk, then an intoxicated defendant is essentially presumed or deemed to be aware of such a risk, even if he is in fact unaware of it. Under the more precise formulation of the Model Penal Code, the intoxicated defendant is considered constructively reckless so long as his factual unawareness of the risk is actually due to his intoxication; but if even a sober actor in the shoes of the defendant would have been unaware, then the defendant is not deemed constructively reckless.<sup>79</sup> Thus,

---

<sup>77</sup> I add "or respond to" those beliefs in order to avoid an overly rationalistic conception of human action. Even if the actor has very little time to think and "consider" his options, he has an obligation to react responsibly to his beliefs about possible harmful results or possible incriminating circumstances.

<sup>78</sup> See TAN 98-150= infra (discussing the cognitive counterfactual approach).

<sup>79</sup> See Model Penal Code, §2.08(2). This special intoxication rule is just one instance of a counterfactual approach to causation, though here, the causal question involves the causal effect of a mental condition (intoxication) on a mental state, not the more typical causal inquiry into whether the actor's conduct caused a result (such as the death of the victim). Ferzan's hostility to counterfactual

the intoxication rule asks a counterfactual question—whether a sober D would have been aware of the risk of which the intoxicated defendant was unaware. (For example, if a drunk driver runs over a pedestrian, he cannot introduce evidence of his intoxication to show that he did not see a pedestrian before hitting him and thus to defeat a charge of recklessness—unless, under the circumstances, even a sober person would not have witnessed the pedestrian.)

This “what if he had been sober?” test is yet a third type of counterfactual criterion relevant to cognitive culpability. The question is not whether, if he had been aware of the risk, the intoxicated actor would still have taken it (the “counterfactual risk-avoidance” test). Nor is the question whether, if he had thought about it for a moment, the intoxicated actor would have been aware of the risk (the “counterfactual consideration” test). However, this third, intoxication test is more closely related to the counterfactual consideration test. For it represents one recurrent type of culpable explanation for not being aware of a risk.<sup>80</sup> And it posits a rough moral equivalence between (a) the culpability of becoming intoxicated and thus blotting out one’s ability to perceive risks, and (b) the culpability of recklessly or consciously taking a risk.<sup>81</sup>

Thus, although the basic connection paradigm for cognitive mental states of knowledge and recklessness is straightforward, many problems arise in applying that paradigm to such contexts as forgetfulness, latent awareness, and intoxication. And a plausible solution to such problems is sometimes the creative use of counterfactual criteria of culpability.

### C. Connecting the actor’s negligence with his acts

Finally, with respect to the mens rea or culpability requirement of negligence, a legally adequate connection to the relevant action is also required.<sup>82</sup>

---

legal criteria extends even to counterfactual tests of causation. Ferzan, *supra* note 9=, at 622-623 n. 87. But this is the prevalent test of factual causation, and it is not clear what test Ferzan would use instead.

<sup>80</sup> However, the intoxication rule differs from the counterfactual consideration criterion in this respect: it does more than constructively elevate latent awareness to conscious awareness when a sober defendant would have had conscious awareness. The intoxication rule also elevates complete ignorance of a risk to conscious awareness; in other words, the rule does not presume that the defendant had at least latent awareness. Thus, suppose a drunk defendant, while driving, is not consciously aware that a pedestrian is near the side of the road. It is possible that, after a collision with the pedestrian, the driver would, if asked, still honestly maintain that he never saw the pedestrian; in other words, the driver lacks latent awareness. But if a sober defendant would have been aware of the pedestrian, then the intoxicated defendant is still treated as if he was aware.

<sup>81</sup> This equivalence is somewhat crude, however. See Glanville Williams, Criminal Law: The General Part 564 (2<sup>nd</sup> ed. 1961); Kyron Huigens, *Virtue and Criminal Negligence*, 1 *Buff. Crim. L. Rev.* 431, 436 (1998); Robinson, *supra* note 52=, at 312-320. For example, “perhaps [the defendant] has got very drunk many times before, and on all previous occasions has fallen peacefully asleep.” Williams, *supra* note 5=, at 103.

<sup>82</sup> Subjectivists might reject any form of criminal negligence liability and insist on a minimum mental state of recklessness (cognitive awareness). Nevertheless, it is instructive to see what connection requirements a negligence standard entails. Moreover, although subjectivists object to negligence liability for a number of reasons, these do not ordinarily include the concern that the requisite connection between defendant’s negligence and his act or omission is lacking.

Consider two different types of negligence—cognitive and conduct.<sup>83</sup> The cognitively negligent actor is one who should have been aware of a risk. (The Model Penal Code employs this conception of negligence.) In this category, the connection requirement is simple. Obviously, an actor’s unreasonable failure to recognize a risk is only legally relevant if this blameworthy omission occurs prior to the relevant action. Negligent homicide requires that the agent, prior to causing death, should have realized that the risk was unreasonably great. A crime of “assaulting a police officer, negligent as to the victim’s status as a police officer” requires that the actor should, prior to the assault, have recognized that the victim was a police officer. In these cases, as in the case of recklessness and knowledge, the requisite connection is implicit: if one had been aware, one could and should have employed that knowledge in declining to engage in the relevant criminal act or to cause the relevant harm.<sup>84</sup>

How do we analyze the second, “conduct” type of negligence? Here, negligence amounts to deficient or grossly deficient conduct (for example, a grossly incompetent performance of a task such as provision of medical care or operation of a motor vehicle). If such conduct is criminalized, what is the connection requirement? Interestingly enough, here the articulation of a connection requirement is not necessary or even possible. For grossly unreasonable conduct just is an act or omission that the actor performs with greatly inadequate skill or care. Since the conduct and the deficiency cannot be disconnected, no distinct connection requirement exists.<sup>85</sup>

To be sure, one can conceive of a free-floating “lack of skill” disconnected from a particular act or omission. A doctor who has badly botched every operation he has performed might have a serious incapacity to provide proper medical care. If he is considering performing another operation but has not yet done so, we might say that his lack of skill has yet to “connect” with a legally relevant act. And it would obviously be improper to punish for such an unexercised potential to act negligently. Still, it is not his general incompetence that justifies punishment when the doctor does perform an operation badly; rather, what justifies punishment is the highly deficient skill revealed in that particular operation. (If, by luck, the doctor performs the procedure as competently as any prudent doctor would, but nevertheless causes harm, legal liability would be unwarranted.) If this is correct, then a free-floating incapacity or incompetence is never relevant to criminal liability.

Thus, the connection requirement for criminal negligence liability is, in the case of cognitive negligence, analogous to the requirement for cognitive recklessness or knowledge. And in the case of conduct negligence, no separate connection requirement exists, because conduct negligence presupposes a direct normative evaluation of the actor’s conduct.

---

<sup>83</sup> See generally Kenneth W. Simons, *Dimensions of Negligence in Criminal and Tort Law*, 3 *Theoretical Inquiries in Law* \_\_ (forthcoming 2002).

<sup>84</sup> Moreover, some of the same temporal proximity problems arise here. Suppose I get in a brief argument with a stranger in a darkened bar, and I should have realized that he was a uniformed police officer. If, at a later date, I start a fight with this person, when he is not in uniform, it does not follow that I then should have realized that he is the same person as the stranger I met in the bar.

<sup>85</sup> In such a case, Antony Duff’s argument that the actor’s state of mind or culpability intrinsically “structures” and gives “meaning” to the action makes sense. Duff, *supra* note 5=, at 162-163. The argument seems inadequate or incomplete, however, for most other culpability terms. See Simons, *Culpability and Retributive Theory*, *supra* note =, at 392; Ferzan, *supra* note 9=, at 611-618.

### III. Connection requirements for culpable indifference

#### A. Introduction: two culpable indifference criteria

Let us now turn to the focal issue. If criminal law employs culpable indifference as a mental state category, it will trigger the criticism that this permits punishment merely for bad character, apart from bad acts. Can the criticism be refuted?

The answer greatly depends on just how “culpable indifference” is understood. And quite a range of different criteria are plausible, and indeed are recognized in legal decisions. This might seem surprising; after all, no similar diversity exists with respect to definitions of purpose, or knowledge, or even conscious recklessness. However, on further reflection this multiplicity of meanings has a straightforward explanation. For it is a natural consequence of a primary role that culpable indifference plays in legal doctrine—namely, the role of serving as a residual (and more explicitly normative) mental state category that supplements, and cannot be reduced to, the conventional categories of purpose, knowledge and cognitive recklessness.<sup>86</sup>

Which version of culpable indifference is most defensible depends on a number of things. For example, does the version articulate a persuasive account of culpability under retributive principles? Is it sufficiently precise to guide and constrain legal decision-makers? But one critical question is whether the version is vulnerable to the “punishment merely for bad character” objection, and that is the question before us. In this section, I will identify two different criteria (or, more accurately, families of criteria) for culpable indifference. The next two sections will examine each criterion more carefully.

The problem of punishing for free-floating mental states is, as we have seen, pervasive. Even conventional mental states of purpose, desire, or belief can be free-floating or disconnected from action in a troublesome way. But culpable indifference seems especially objectionable. The very term “indifference” connotes a mere attitude of not caring about the consequences of one’s actions. Accordingly, on first impression, it is difficult to see how the first two examples

---

<sup>86</sup> Depraved heart or extreme indifference murder, for example, is a category meant to include killers who are as dangerous or as blameworthy as those who intend to kill or who knowingly kill. For further discussion, see TAN 171= *infra*.

Variants of culpable indifference can also serve as an aggravating factor in homicide, helping to justify the imposition of the death penalty. Thus, in *Arave v. Creech*, 507 U.S. 463 (1993), the Supreme Court upheld against a vagueness challenge the aggravating circumstance that the killing was done “in utter disregard for human life.” The Court found that the state court’s “narrowing” construction of this language—that the crime must “exhibit the highest, the utmost, callous disregard for human life, i.e., the cold-blooded, pitiless slayer”—was constitutionally adequate. In this context, a kind of “extremely extreme indifference” test is employed to help select those first-degree murders (including premeditated intentional killings) that deserve the death penalty. In this article, by contrast, I focus on the use of culpable indifference as a mental state for an offense comparable (for grading purposes) to purpose, knowledge, and recklessness. However, the vagueness and connection problems that this paper addresses obviously should be of great concern in the death penalty context as well. For example, the Court in *Arave* viewed the narrowing construction as limiting application of the circumstance to those who kill “without feeling or sympathy.” *Id.* at 476. It is doubtful that this standard satisfies the connection requirements I have discussed, for the lack of feeling or sympathy apparently could merely amount to a (disconnected) lack of remorse.

in the introduction differ from the four examples that followed; in each case, it appears, a criminal sanction would simply punish a person for revealing a bad attitude, and would thus be indefensible.

In order to avoid this objection, culpable indifference, like any other mental state or culpability requirement, must be appropriately expressed in, or connected to, action. Recall introductory examples C, D, E, and F. In each example, it is intuitively clear that the actor has not merely expressed a passive attitude of indifference. The torturer who leaves the victim to die, and the Russian roulette player who willingly imposes a 1/6 chance of death for a personal thrill, each displays grossly insufficient concern for the interests of their victims. (In this most aggravated form, culpable indifference can approximate the culpability of intentional wrongdoing.) Moreover, the indifferent speeding driver who believes he is posing little or no risk to others might be as culpable as a driver who is aware of a significant risk yet makes some effort to avoid it. And the indifferent statutory rapist who gives no thought to the age of the victim might be as culpable as the individual who realizes that there is some risk that she is underage but (negligently) concludes that the risk is trivial. (In this more modest, less aggravated form, culpable indifference might approximate the culpability of cognitive recklessness.)

In each of these cases, the actor's culpable indifference or lack of concern for the interests of others is, in some sense, expressed in his action. But what counts as "expression in action"? What is the connection requirement for culpable indifference?<sup>87</sup> Consider two possible answers, premised on two plausible (but rather different) counterfactual criteria.

#### *1. Idealized criterion.*

One possibility is a direct counterfactual criterion, calibrated to the conduct of an idealized, non-indifferent person. Under such a counterfactual ideal person criterion, an actor's modest indifference is connected to his action just in case he would not have taken the action in question if he had shown socially appropriate concern for others.<sup>88</sup> (If the speeding driver had cared sufficiently about the interests of his potential victims, he would have maintained a lower speed and made efforts to avoid injury. And if the statutory rapist had cared about the age of his victim, he would have given some thought to her age, would have inquired about it, and then would have discovered that she was underage.)

Similarly, under this criterion an actor's extremely indifference is connected to his action just in case the action he took was an extreme departure from the conduct we would expect of a non-indifferent person.<sup>89</sup> Under this

---

<sup>87</sup> For my earlier thoughts on this issue, see Simons, Culpability and Retributive Theory, at 391-393.

<sup>88</sup> An alternative formulation is this: modest indifference is connected to action if, in acting, the actor gave less weight to the interests of others than he should have. However, framing the question this way—what is the "optimal weight" in balancing one's interests against the interests of others?—is problematic, as Kim Ferzan has pointed out. Ferzan, *supra* note 9, at 626 n. 96. The degree of moral depravity expressed in an act is not properly understood simply as a matter of the discordance between the weight the actor gave to the interests of others and the socially optimal weight that those interests deserve. "Proper" and "improper" weight cannot be determined by a simple context-independent formula.

<sup>89</sup> Moreover, arguably even the modestly indifferent actor should be defined, not as one who departs from the ideal standard to any degree, but only as one who departs rather significantly. See Simons,

criterion, the torturer (C) who harms a victim and does not care whether the victim dies is even more culpable than a mercy-killer, even though the latter intends to kill, because the mercy-killer at least is motivated by a desire to relieve the victim's suffering. Accordingly, his indifference to the interests of the victim is less. And similarly, the Russian roulette player (D) who willingly endangers another's life for the sake of a personal thrill is extremely indifferent to the victim's fate.

Moreover, the acts of the indifferent speeding driver (E) show a serious degree of culpability, more serious than the culpability of a different speeding driver (E2) who creates the same degree of risk of injury but only after trying to minimize any harm to others. And similarly, the statutory rapist (F) who demonstrably has no interest in whether the victim is below the age of consent displays more culpability than an actor (F2) who unreasonably concludes, after good faith inquiry, that the victim is not below the age of consent.

Notice that the direct counterfactual criterion is similar to a negligence test insofar as it directly measures conduct against an idealized normative standard. And, like a negligence test, the direct counterfactual criterion resolves the "connection" problem in a very simple way—by dissolving it!<sup>90</sup> For one cannot coherently say both that an act is deficient relative to an idealized standard (whether of non-negligent conduct, or of non-indifferent conduct) and that the culpable deficiency is "disconnected" from the act.<sup>91</sup> Thus, in all four cases, the connection requirement poses no additional difficulty, once we have indeed determined that the acts of the defendant are modestly or extremely deficient relative to the idealized standard. A genuine difficulty does arise, however, in specifying clearly how one should make that determination, as we will see.

## 2. *Cognitive counterfactual criterion.*

A second possible culpable indifference standard also employs a counterfactual criterion, but of a very different kind: it asks what the actor would have done if he had had a different belief about the relevant risks. (Earlier, I described this as the counterfactual risk-avoidance test.<sup>92</sup>) Under this cognitive counterfactual approach, the modestly indifferent actor is one who was unaware of a risk but who would have confronted it even if he had been aware of it, i.e.,

---

Culpability and Retributive theory, *supra* note 5=, at 386-387. Otherwise, criminal law will punish actors of very modest blameworthiness. (Most cases of ordinary tort negligence could be characterized as reflecting deficient concern for the interests of others, but not all such cases deserve criminal punishment.)

<sup>90</sup> See TAN 82-85= *supra*.

<sup>91</sup> To be sure, one's culpably indifferent act or course of conduct can create a risk, without the risk taking fruition in harm, just as one's negligent, reckless, knowing, or purposeful act can fail to cause harm. (Examples would be an actor's dangerous, indifferent driving; or his playing Russian roulette and pulling the trigger once, with the bullet fortuitously absent from the firing chamber.) But the possibility of the act itself not causing the harm does not, in the relevant sense, create a "disconnection" problem. Rather, that problem arises when the mental state of the actor is not appropriately linked to the act or the omission that could cause the harm.

<sup>92</sup> See TAN 75= *supra*. The counterfactual consideration test is also a counterfactual cognitive test. See TAN 75= *supra*. But it applies in more limited circumstances and thus does not receive further discussion here.

even if he had been (cognitively) reckless.<sup>93</sup> The extremely indifferent actor is one who did not realize that the risk was likely to occur (i.e., he was not “knowing”) but who would have confronted it even if he had realized that it was likely.<sup>94</sup> Under this criterion, roughly speaking, modest culpable indifference combined with negligent unawareness is punished as harshly as cognitive recklessness, and extreme culpable indifference combined with cognitive recklessness is punished as harshly as knowledge.

How would this test resolve the introductory examples? The torturer is extremely indifferent and equivalent in blameworthiness to one who knowingly kills if (as might well be the case) he would have persisted in his conduct even if he had known that it would cause death. The Russian Roulette player might also be extremely indifferent under this criterion: again, it might be the case that he would have shot the victim even if he had known that death was very likely to follow. (However, this is much less clear than in the torturer case, as I will clarify below.) The indifferent speeding driver also would be extremely indifferent if he would have continued driving in the same manner even if he had known that injury would occur. By contrast, the statutory rapist is modestly indifferent if, although he was unaware of any risk that the victim was under age, he would have proceeded anyway even if he had suspected that she might be.

Under either the idealized or the cognitive criterion, however, the first two examples from the introduction clearly fail the connection requirement. The September 11 celebrants satisfy neither criterion, for they have engaged in no criminal action at all, but have merely expressed approval of the criminal actions of others.<sup>95</sup> The remorseless negligent driver has indeed caused a serious harm, for which he might be criminally responsible. But his lack of remorse has no connection to his prior risky conduct. Applying the idealized criterion, we have no reason to believe that his subsequent attitude of indifference had anything to do with his distraction while driving. (Contrast this case: a driver tells his passenger that he could care less whether he causes an accident, and then puts on a blindfold, resulting in a deadly accident.) And, applying the cognitive criterion, the driver’s lack of remorse provides only very faint evidence that, had he been aware of the risk *ex ante*, he would have taken it. (Again, contrast the driver who puts on a blindfold.)

However, the Russian Roulette example illustrates an important difference between the two counterfactual criteria. In the example, the defendant places only one bullet in the gun, spins the barrel, and credibly commits to shooting only once. The cognitive criterion is unlikely to treat the defendant as a murderer, for we have no reason to believe that the actor would play the game even if he believed that the probability of death was a practical certainty. (Indeed, to the contrary, we have good reason to believe that he would not play under those circumstances; presumably the thrill of the game flows at least in

---

<sup>93</sup> Moreover, the even less culpable, mildly indifferent actor could be defined as one who would have been aware of the risk if he had cared more; but who then would have declined to take the risk. Jeremy Horder calls such an actor “weakly” indifferent, in comparison to a “strongly” indifferent actor. Horder, *supra* note 5=, at 502-3. (The latter is identical to the “modestly indifferent” actor I identify in the text.)

<sup>94</sup> Alan Michaels endorses this latter conception under the label “acceptance.” Michaels, *supra* note 5=. I discuss his analysis below.

<sup>95</sup> Nor have they given encouragement in advance of, or contemporaneous with, the criminal acts of others. When such encouragement facilitates a crime, accomplice liability is appropriate.

part from uncertainty about whether death will result.<sup>96</sup>) But the idealized criterion can condemn this actor's conduct as extremely culpable: imposing a substantial risk of death for a malicious and clearly unjustifiable reason arguably amounts to "extreme indifference."

A similar analysis applies in another paradigmatic depraved heart murder case—where an actor fires a gun into an occupied house in order to scare the inhabitants. This case might fail the cognitive criterion, for the actor might not have been willing to shoot if he was practically certain that he would kill. However, his action might satisfy the idealized criterion. Finally, a grisly variation of the torture example also illustrates the difference between the two criteria. Suppose that an especially malicious torturer wants to keep the victim alive in order to torture him again; he wants to cause as much pain as possible but keep the victim suffering. The victim nevertheless dies. Such a torturer would not be extremely indifferent under the cognitive test: if he knew that a certain type or level of torture would cause the victim's death, he would most certainly avoid it. Nevertheless, there is much to be said for the idealized criterion's judgment that in this case, too, the torturer is extremely depraved and deserves a punishment as severe as that deserved by most knowing or intentional killers.<sup>97</sup>

Each criterion has distinctive advantages and disadvantages. Generally speaking, the cognitive criterion has the advantage of greater clarity, precision, and, to some extent, predictability. But it also seems unduly narrow in some cases, such as the examples just discussed. On the other hand, while the idealized criterion applies to a more comprehensive range of culpable acts, it has the significant disadvantage of employing an extremely vague and often indeterminate criterion.

The next two sections evaluate each criterion with greater care.

#### B. The cognitive counterfactual criterion: further analysis

Professor Alan Michaels has provided the most careful and thorough analysis of what I call the cognitive counterfactual criterion. He uses the term "acceptance" to describe the mental state of an actor who (a) is reckless<sup>98</sup> as to a circumstance or result (i.e., is aware of a substantial risk that the circumstance exists or that the result will occur), but (b) would have taken the risk even if he knew that the circumstance was likely to exist or that the result was likely to occur. An "accepting" actor, he argues, should receive the same punishment as a

---

<sup>96</sup> Contrast a case in which the defendant plays this game with six involuntary victims, and commits to pointing the gun and pulling the trigger once at each victim until someone dies. In this case, if the first victim happens to die, then although the ex ante chance of death was only one in six, the cognitive criterion is clearly satisfied.

<sup>97</sup> Similarly, with respect to circumstances rather than results, the cognitive criterion again might be too narrow. In the context of wilful blindness, for example, perhaps an actor who deliberately avoids knowledge should be deemed to "know" that a circumstance exists even if he would not have acted had he known. See Kim Ferzan, "Contribution to this volume," n. 101 (posing the example of a wilfully blind actor who flunks the counterfactual test yet does everything in his power not to know).

<sup>98</sup> Michaels defines recklessness in the Model Penal Code's sense (Michaels, *supra* note 5, at 963-964), but indicates that his approach can conform to whatever definition of recklessness the jurisdiction happens to employ. *Id.* at 961 n. 23. However, I believe that this proviso presumes that recklessness at least requires cognitive awareness of risk; the acceptance approach could not readily apply if recklessness required only gross negligence.

knowing actor.<sup>99</sup> One could also expand this counterfactual analysis to the culpably unaware (or “cognitively negligent”) actor: such an actor might be punished the same as a cognitively reckless actor if he (a) is negligently unaware as to a circumstance or result, but (b) would have taken the risk even if he was (recklessly) aware of a substantial risk that the circumstance existed or that the result will occur.<sup>100</sup> (Discussion of this extension of the “acceptance” approach is postponed until later in this section.)

Two of Michaels’ examples offer powerful intuitive support for his “acceptance” test. The first concerns results.

#### H. Slumlord

Slumlord decides to burn down one of his apartment buildings to collect the insurance. He believes that the fire is almost certain to burn down the adjoining building, too. As he is about to set off a remote control incendiary device to start the fire, Slumlord receives a phone call revealing that the adjoining building contains a new sprinkler system that will greatly reduce the risk of the building burning down if a fire starts. Accordingly, when Slumlord presses the remote control, he is only reckless as to the risk of the adjoining building burning down, and no longer believes that that result is almost certain. Nevertheless, the adjoining building does burn down.

The culpability of Slumlord for burning down the adjoining building is the same, whether he pressed the remote control before or after receiving the phone call. In either case, he “accepts” the result that occurs.<sup>101</sup>

A second example from Michaels concerns acceptance of a circumstance, rather than a result:

#### I. Sleaze

Mr. Sleaze needs a topless dancer for his bar. It is illegal to employ a person under the age of eighteen as a topless dancer. Mr. Sleaze auditions four people for the job. He thinks one of them, Ms. Minor, is great and hires her. Later, Mr. Bouncer tells Mr. Sleaze that he checked the I.D.'s of the four people who auditioned, and two of them clearly had altered dates of birth. Therefore, two dancers must be under age, but Bouncer does not remember which two. Sleaze says, "I don't care whether Ms. Minor was under age or not--she's going to make me money," and never checks Ms. Minor's age further. In fact, Sleaze would have hired Minor even if she had told him she was seventeen. If she is

---

<sup>99</sup> With respect to circumstances, Michaels believes that his acceptance approach gives the most justifiable explanation of the doctrine of wilful blindness, under which reckless actors are sometimes treated as possessing knowledge if they are culpable in not obtaining knowledge (for example, if they deliberately blind themselves to the facts). With respect to results, he believes that his approach gives the most justifiable explanation of depraved heart murder.

<sup>100</sup> Michaels does not extend his analysis this far, however. *Id.* at 962 n. 26.

<sup>101</sup> I have paraphrased and simplified Michaels’ original example, *id.* at 969-970.

underage, the high probability [definition of knowledge] would say that Sleaze lacked the culpability of knowledge (assuming fifty percent is not a high probability). [Instead, Sleaze would only be reckless.] On the other hand, if Bouncer said three of the identifications were altered, the result might change (if seventy percent is a high probability), even though it makes absolutely no difference to Sleaze.<sup>102</sup>

[In either case, Michaels concludes, Sleaze accepts the unlawful circumstance (hiring an underage person).]

Appealing as these two examples are, however, sometimes counterfactual culpability is highly problematic. Recall Larry Alexander's "Deborah" example (G), above:

#### G. Deborah

Suppose ... that Deborah is engaging in her Sunday driving within the speed limit and believes she is creating a low level of risk to others. And suppose that the pleasures of Sunday driving are sufficient to justify imposing that low level of risk on others. But suppose also that Deborah is callous towards the risks she imposes on others, and that even if she were to come to believe that the risks imposed by her Sunday driving were much higher than she actually believes them to be, she would continue to drive notwithstanding that belief. On my account and the law's, Deborah is not acting recklessly at this moment and is therefore not culpable. She has an undesirable character trait, one that might dispose her to act recklessly if the circumstances were right, but her counterfactual culpability does not translate into present culpability.<sup>103</sup>

In the remainder of this section, I explore in detail: (a) some problems that very broad counterfactual mental state tests can pose; (b) six implicit conditions of Michaels' acceptance test that explain why the Slumlord and Sleaze examples are persuasive; (c) problems with extending the acceptance approach to actors who are culpably ignorant, or even nonculpably ignorant, of a risk; (d) the reasons why neither Ferzan's "Jane" nor Alexander's "Deborah" exemplify a fatal defect in the acceptance approach.

Let us begin with the broad question whether the law ever does or should employ counterfactual criteria of culpability. The discussion above reveals that in a number of circumstances, such criteria are legitimately employed. Thus, the mental state of "purpose" is helpfully analyzed by means of a counterfactual test: in hypothetical and relevantly different circumstances, would the actor continue to take actions to bring about the legally proscribed harm, or instead fail to take advantage of such opportunities? Would his further purposes be satisfied even if the harm did not occur?<sup>104</sup> With respect to latent knowledge, one might ask whether, if the actor had considered the matter for even a moment, he would have acknowledged his belief.<sup>105</sup> Similarly, rules about intoxication often consider whether, if the actor had been sober, he would have been aware of the risk of which the intoxicated actor was unaware; if the answer is no, then the

---

<sup>102</sup> Michaels, *id.* at 983.

<sup>103</sup> Alexander, *supra* note 9=, at 938. (This excerpt omits the last paragraph.)

intoxication did not actually cause the actor's loss of awareness, and the constructive recklessness approach to intoxication is then unwarranted.<sup>106</sup>

Of course, any "objective" criterion of criminal liability appeals to an idealized standard, either of acts or of beliefs, and often the standard is the reasonable person. (Indeed, below I analyze in more detail an idealized standard of culpable indifference.) In a limited sense, such a standard is counterfactual:<sup>107</sup> the question is whether the actor would have acted the same way if she had had the qualities of the idealized person. However, in employing such a test, Kim Ferzan points out, "we are judging the defendant against a standard, not punishing him based on the prediction of his actions under different circumstances."<sup>108</sup> The counterfactual nature of such a test is therefore not especially problematic.

On the other hand, reconsider Alexander's assertion that punishing Deborah is just one instance of punishing for a character trait "that at most makes [one] a counterfactual wrongdoer." Depending on its breadth, this objection can indeed be valid. It can be problematic to premise an actor's criminal punishment on what he would have done, had his mental state been different, rather than on what he actually did, given his actual mental state. Consider some examples of counterfactual criteria that are clearly objectionable:

(1) Antonio very likely would have committed a particular crime if physical circumstances had been different (if a ten dollar bill had been left on the sidewalk on a deserted street, if his angry son had provoked him in a particular manner, and so on). This hardly justifies punishing him when he did not engage in the relevant acts at all.<sup>109</sup> A very broad "counterfactual culpability" approach might justify punishment here: we might be able to predict that, if the actor had faced a particular temptation, he would commit a criminal act with the requisite culpability.

(2) Betty very likely would have committed a crime if she had had a different intention (for example, she would have killed someone if she had formed the firm intention to do so). This hardly justifies any criminal punishment if she has not in fact formed that intention.

(3) An inventive example from Kim Ferzan:

---

<sup>104</sup> See TAN 47-48= supra.

<sup>105</sup> See TAN 75= supra.

<sup>106</sup> See TAN 79-81= supra. Another example of the criminal law's use of a counterfactual test is the entrapment defense, which often requires proof that the defendant was "predisposed" to commit the offense, in the sense that the defendant would have committed the crime even if the government had not overreached or had not been involved. See Michaels, supra note 5=, at 1021-24.

<sup>107</sup> One might object that like any counterfactual test, the idealized counterfactual test is inherently problematic. But the counterfactual nature of an idealized reasonable person or non-indifferent person isn't itself problematic for, as explained above, there is no distinct "connection" requirement for such tests. TAN 85= supra.

<sup>108</sup> Ferzan, supra note 9=, at 622 n. 87 (critiquing Michaels' argument that the law's widespread use of "counterfactual" reasonable person tests supports the use of the counterfactual "acceptance" approach).

<sup>109</sup> See Thomas Nagel, Moral Luck, Mortal Questions 28 (1979) (discussing "circumstantial luck").

J. Jane

[I]magine Jane who is driving her car, and fails to notice John in the crosswalk because she is too self-absorbed. Now, Simons would say that if Jane would have continued on anyway, had she in fact seen John, she should be punished as culpably indifferent, rather than simply negligent. But let us assume a little history between Jane and John. Jane, in fact, hates John. Indeed, Jane was driving to John's house, where she planned to take out her shotgun and murder John. Now, if we are playing the "what if" game, why stop at culpable indifference? Isn't it in fact more likely that Jane would aim at John and purposefully (not culpably indifferently) run him over?<sup>110</sup>

Our unwillingness to punish in these three cases rests in part on the view that the state should not punish for "character" alone. Predicting the criminal acts that Anthony would likely commit if he were tempted raises this concern quite directly; he has not even engaged in any preliminary steps towards effectuating his predicted intention. Betty has not even displayed a dangerous or culpable character; any of us might commit murder if we were to form the firm intention to do so! Jane is a more complex case (and is discussed further below).<sup>111</sup> Perhaps her acts prior to the collision suffice for an attempted murder conviction.<sup>112</sup> But she certainly is not liable for murder, even if we can predict that, had she seen John prior to the collision, she likely would have killed him intentionally rather than accidentally.

Why, then, are the Slumlord and Sleaze examples intuitively compelling? Why, in these circumstances, do "accepting" and knowing actors seem to be equally culpable?

A first, initially plausible argument is this. Perhaps the cognitive counterfactual or acceptance criterion follows from a basic grading rationale for distinguishing reckless from negligent actors.<sup>113</sup> Recall the following reason for treating knowing actors more harshly than reckless actors: the former have chosen to act despite a known very high risk of harm, while the latter (who believe the risk is significant but not necessarily very high) might not have chosen to create the risk if they had believed it to be very high. It might then follow (as the acceptance approach provides) that if the reckless actor would have acted the same way even if he had believed that the risk was very high, he should be treated the same as the knowing actor.<sup>114</sup>

Alas, the argument does not strictly follow. The grading rationale justifies different punishment of reckless and knowing actors, given the different

---

<sup>110</sup> Ferzan, *supra* note 9=, at 622-623.

<sup>111</sup> See TAN 135-140= *infra*.

<sup>112</sup> Under the broad Model Penal Code "substantial step" test, the sufficiency of her actus reus would likely be a jury question. But Ferzan rejects attempt liability here and in any other case in which the attempt is not "complete," i.e., in which the defendant has not taken the last act she believes necessary to complete the crime. See Larry Alexander & Kimberly D. Kessler, *Mens Rea and Inchoate Crimes*, 87 *J. Crim. L. & Criminol.* 1138, 1168-1174 (1997). "Kessler" is Ferzan's former name.

<sup>113</sup> See TAN 77= *supra* (discussing a "negative counterfactual rationale" for the grading difference).

<sup>114</sup> An analogous argument could justify an acceptance or cognitive counterfactual approach as applied to some unaware actors: if they would have acted the same way even if aware of the risk, then arguably they should be treated the same as the recklessly aware actor.

types of risks they actually comprehended and thus the different kinds of choices they made. But the acceptance (or counterfactual equivalence) rationale justifies equivalent punishment of knowing actors and some reckless actors—namely, those reckless actors who would have acted the same way if they had known that the risk was likely. In other contexts, simply because it is justifiable or proper to impose lesser punishment on (1) a given act with actual mental state A than on (2) the same act with (different and more culpable) actual mental state B, it clearly does follow that it is justifiable or proper to impose equivalent punishment on (2) an act with actual mental state B, and (3) the same act with actual mental state A but counterfactual mental state B. Consider, for example, the differential between purpose and recklessness. Although it is justifiable to punish actors who recklessly cause harm less harshly than actors who purposely cause harm, it hardly follows that we should provide equivalent punishment to the latter actors and to those reckless actors who would have caused the harm if they had had the purpose to do so—although they lacked such a purpose!

But there are a number of other features of the Slumlord and Sleaze examples that explain their compelling force.<sup>115</sup> The first four features in the following list are especially important in explaining the persuasive force of the examples. (The third, we will see, applies to Slumlord but not to Sleaze.)

(1) The only difference between the actor's actual mental state at the time of the crime and his counterfactual mental state is a difference in the magnitude of the risk that the actor subjectively perceives.

(2) We can confidently predict that the actor would take precisely the same action, whether he had the actual mental state (reckless) or the counterfactual one (knowledge).

(3) The basis of this prediction is that when the actor was initially prepared to take the action, he possessed the "higher" mental state of knowledge, but by the time he acted, his mental state had "deflated" to reckless.

(4) Very little time elapses between (a) the moment when the actor reveals a willingness to take the action with a belief that the risk is very likely, and (b) the moment when he takes the action with the (merely reckless) belief that the risk is less likely.

(5) The change in the actor's subjective belief about the magnitude of the risk does not result from the actor's own conduct.

(6) The prediction that the reckless actor would have acted the same way even if he had known that the risk was likely rests on an appropriate causal connection. In rare cases, that is, the prediction might

---

<sup>115</sup> Michaels provides a general argument for equating the culpability of knowledge with the culpability of acceptance (even if the actor lacks knowledge). From a retributive perspective, the merely accepting actor still shows the same culpable attitude towards the result or circumstance as the knowing actor shows: a willingness to permit the known harm in pursuit of his own ends. Michaels, *supra* note 5=, at 966-970. (Michaels' argument that a utilitarian perspective also supports equating the two mental states, *id.* at 971-976, is beyond the scope of this article.)

The account of relevant factors that I offer in the text below is an attempt to explain more precisely why and when acceptance satisfies retributive principles. However, this account probably differs from Michaels' in some respects, especially in its rejection of the pure "inflationary" scenario, discussed below.

be true only because of an unusual sequence of events that does not reflect equivalent culpability between knowing and “accepting” actors.

Consider these features in turn.

Factor (1) is crucial. Counterfactual mental states of belief are not nearly as problematic as counterfactual mental states of intention or purpose. The fact that Betty (above) would have killed someone intentionally if she had had such an intention is obviously irrelevant to her criminal culpability. It is irrelevant even if she happens to negligently or faultlessly cause a death, and even if she would (consistent with factor (2)) have taken precisely the same action if she had formed the intention to kill.<sup>116</sup> Or suppose, to draw a closer analogy to the comparison of cognitive states (reckless and knowing), we can predict that an actor who actually acts with a “bad” intention would have acted the same way if (counterfactually) he had had an even “worse” intention. Hector punches Victor in the nose, intending to cause a slight injury but not to kill him. If he had wanted to kill Victor, let us suppose, Hector would have punched him in exactly the same way—as forcefully as he can, which it turns out is not very forceful. (And suppose more deadly means were not available.) It hardly follows that Hector should be punished as harshly as one who intends to kill, when he never had that intention.<sup>117</sup>

Why the asymmetry? For the following reason: when one’s subjective beliefs about the magnitude of a risk change, they normally change for reasons having nothing to do with one’s culpability, but a change in one’s intentions of the sort just noted clearly does reveal greater culpability.<sup>118</sup>

Factor (2) underscores the very limited scope of the counterfactual prediction that the acceptance test calls for. Slumlord, Sleaze, and the other “acceptance” examples involve an actor who engages in (or would engage in) the

---

<sup>116</sup>Suppose the relevant action is this: she drives through an intersection at a particular speed, and the victim, a pedestrian, is directly in front of her. Such an action could be performed intentionally, negligently, or faultlessly. Or consider an even more extreme version of Ferzan’s Jane example, noted above, that resembles Slumlord. As before, Jane performs the action of driving her car at a given speed through a crosswalk, an action that accidentally causes John’s death. That morning, she had formed the intention to kill John. But by the time she drives to his house, she no longer has that intention (just as Slumlord, by the time he presses the button, no longer has the counterfactual mental state of knowledge). Obviously it is irrelevant that she once harbored an intention to kill, if she no longer has that intention at the time of the killing and has taken no steps to effectuate that intention.

<sup>117</sup> Even if Hector initially had the intention to kill, but then changed his mind and decided only to injure slightly, his earlier intention is irrelevant to his responsibility (assuming he did not take any action to execute the earlier intention). Proof that he had indeed changed his mind might be difficult, of course, since the external act (punching as hard as he can) is by hypothesis the same. But in principle, the connection requirement for intentions precludes criminal liability for a prior intention that no longer motivates the act in question.

<sup>118</sup> For further discussion of this asymmetry, see Simons, *Culpability and Retributive Theory*, supra note 5=, at 376-377. An exception, I believe, is where the change in intentions is due to a morally exculpatory cause, such as duress or severe mental illness.

Michaels might analyze factor (1) differently—by rejecting its premise. For he concludes that the same actual mental state, “acceptance,” is present in both cases. Actual knowledge and counterfactual knowledge (plus recklessness) are simply two different ways of exhibiting an “accepting” state of mind. See Michaels, supra note 9=, at 967-968. However, this analysis does not really answer the objection that punishing the accepting actor is problematic when such acceptance is premised on a counterfactual belief state.

same risk-creating conduct whether his mental state is “lower” (recklessness) or “higher” (knowledge). The actor is not being counterfactually punished for a different act that we predict he would have committed (if his mental state had been different), nor for potentially causing a greater harm than he actually brought about. Thus, if Slumlord initially considered using a much more dangerous incendiary device that he believed was likely to kill several people in the neighborhood, and then decided to use the less dangerous device that he believed would simply burn two unoccupied buildings, employing a counterfactual test to punish him as harshly as one who knowingly kills would be very troublesome.<sup>119</sup>

The predictive test must be limited in this manner to ensure that it only encompasses an actor whose actions clearly reveal his culpability. Expanding the prediction to different acts is qualitatively different. If an actor decides to proceed with a different act that he believes to be less dangerous, we usually have little basis for assuming that he remains willing to commit the initially contemplated, more dangerous act. But such an assumption is often plausible if the actor’s conduct does not change at all and if, instead, only his beliefs change, for reasons unrelated to his willingness to commit the crime.<sup>120</sup>

Under Factor (3), mental state “deflation” is a more persuasive basis than mental state “inflation” for a counterfactual prediction of culpability. In Slumlord (though not in Sleaze), the mental state “deflated” from knowledge to recklessness. That is, the slumlord initially chose a course of action with the belief that he was likely to cause harm; but immediately thereafter, he came to believe that the risk was only “substantial” (thus, he was only reckless, not knowing); yet he continued with his original plan.

In such a case, the actor, when he initially decides to create (what he believes to be) a very high risk, gives much less weight to the harmful side-effect than he should have. This decision is immediately followed by a change in information suggesting to the actor that the risk is not so high. But the circumstances relevant to criminal liability do not change in any other way. Accordingly, we have good reason to conclude that the actor’s attitude to the risk has not changed: he would still be willing to create the higher level risk, if the information were to change once again (“inflate”) and cause him to believe once again that the probability is high.

---

<sup>119</sup> To some extent, of course, attempt liability does permit punishment for harms the actor was “willing” to produce. However, attempt liability also requires proof that the actor took steps beyond preparation towards producing such a harm, and many jurisdictions further recognize a renunciation defense.

<sup>120</sup> Of course, the action that must be the “same” for purposes of this criterion is the action that creates the risk of harm, and not (for example) Slumlord’s action of picking up the phone to receive the call informing him that the risk has lessened, nor a willfully blind actor’s conduct in deliberately avoiding knowledge of a risk. The fact that the latter actor would not bother to avoid inspecting an item if he already had knowledge (and in that sense would have taken a slightly different action if he were knowing rather than reckless) is irrelevant. I thank Alan Michaels for clarifying this point.

What if the knowing actor decides to employ a less dangerous method for nonexculpatory reasons, reasons that suggest he is still just as willing to bring about the harm? (Suppose Slumlord decides to use a less dangerous device, one that lowers the risk of death from knowing to reckless, only because this device is more difficult for the authorities to trace to him.) Here, it would be plausible to relax the “same action” requirement so that such an actor could be characterized as “accepting.” See also Michaels, *supra* note 5, at 997 n. 154, 1002 n. 170. However, defining this exception clearly and applying it accurately might be sufficiently difficult as not to be worth the trouble.

Contrast Michaels' "deflationary" Slumlord scenario with the following "inflationary" scenarios:

H-2. Slumlord: "inflation" variation

Suppose Slumlord initially believes that there is only a small risk that the adjoining building will burn down, because he believes that a new sprinkler system has been installed that would minimize the risk of a serious fire. He presses the remote control and causes the fire. Then he is told that actually, the new system was not yet functioning. Accordingly, he now believes that there is a very high risk that the building will burn down. But he states, "Well, I would have started the fire anyway."

I. Sleaze [abbreviated from above]

Recall that Sleaze is initially informed of a two in four chance that Ms. Minor is under age. He indicates that he doesn't care, but wishes to hire Minor without regard to her age. He hires her. His comment suggests that even if he were later informed that the chance that she is under age was three or even four out of four, he would have hired her.

In both of these cases, we have a new reason, not applicable to the deflationary scenario, not to punish the actor despite his counterfactual culpability of knowledge. That new reason is the principle of respecting the actor's autonomy. Specifically, the actor should be free to change his mind, even if at one point in time he firmly intends to commit a serious crime. Moreover, an actor should also be free to entertain immoral thoughts, beliefs, or intentions if he does not genuinely express them in his actions.

When an actor, at the time he caused a harm, was only reckless as to the harm occurring, and when the actor reveals his willingness to take even a higher risk of harm only after his decision to act, it is especially dangerous to let punishment turn on such an ex post judgment. In this "inflationary" scenario, unlike the deflationary scenario, the actor did not express acceptance (a willingness to create a very high risk of harm) at any time prior to his act. An expression of acceptance after performing the criminal act usually permits a less reliable prediction (or, more precisely, a less reliable ex post estimate) of what the actor would have done. But, more importantly, even if the estimate is empirically just as reliable, relying on such an estimate is inconsistent with a proper respect for the actor's liberty. Unless the actor has actually expressed a willingness to cause the harm or engage in the criminal act knowingly, at a point in time when causing the harm or engaging in the act is a decision actually open to the actor, it is unjustifiable to rely only on evidence that later, the actor expressed a willingness so to act if the occasion were ever to repeat itself.<sup>121</sup> The

---

<sup>121</sup> "[T]he act requirement provides a locus poenitentiae to enable the law-abiding citizen to avoid criminal liability." John Jeffries & Paul Stephan, *Defenses, Presumptions, and Burden of Proof in the Criminal Law*, 88 Yale L. J. 1325, 1371 n. 130 (1979), referring to Herbert Packer, *The Limits of the Criminal Sanction* 73-75 (1968). The policy of not relying even on accurate predictions is strongest in the attempt context, where the actor has not even committed the actus reus of the offense. In the

acceptance test should accordingly be restricted to what I have called the “deflationary” scenario—with a qualification noted below.

The earlier discussion of the connection requirement for knowledge<sup>122</sup> is instructive here. Recall the example of an actor who acquires knowledge prior to committing a criminal act, but then, when he acts, forgets what he knows or is distracted from employing his knowledge. Such an actor might still be liable for a crime requiring knowledge, if we conclude that he has a continuing duty to employ his latent or earlier-acquired knowledge. (Suppose someone loads his gun in the morning and forgets that it is loaded when he carries it in public in the afternoon.) This continuing duty is especially defensible if we are confident that he would have acted the same way if he had had knowledge at the time of his act. (Suppose that in the morning, when the actor loaded his gun at home, he planned to transport it, still loaded, in public that afternoon). On the other hand, when the concurrence problem corresponds to the “inflationary” scenario, the continuing duty argument is inapt, just as the analogous acceptance argument is much less defensible. (If the actor plans to transport (what he believes to be) an unloaded gun in public, and only after the act of carrying the gun in public acquires a belief that it is loaded, the duty to employ that knowledge in his action most certainly cannot “relate back” to the original act. And this is so even if we predict that he would have carried the gun in public if he had been aware it was loaded.)

Now the qualification: the concern we should have about adopting an acceptance approach in the inflation scenario is much less when the actor expresses his “inflationary” attitude prior to committing the criminal act. Thus, suppose that Slumlord, in the last variation (H-2), expresses his willingness to create the risk even if it was much higher, prior to pressing the remote control button and igniting the fire. That is, he is informed that the risk is probably likely, not merely substantial; he still believes that it is merely substantial; but he indicates that in any event, that difference in probability would not affect his decision. In this situation, the autonomy values noted above are not as seriously implicated. For we are no longer taking a post-act attitude and permitting it to

---

present context, the actor has indeed committed the actus reus, and indeed has done so with a culpable (reckless) state of mind. Nevertheless, it remains troubling to impose greater punishment based on a prediction that he would still have committed this actus reus if he had had a more culpable mental state, even though he never possessed that mental state until after he committed the actus reus.

To be sure, even the deflation version of acceptance is in tension with autonomy values, insofar as the actor might, at the later moment when he recklessly acted, no longer have been willing to knowingly cause a harm. But the actor’s willingness to create a known high probability harm when that choice was actually available makes this version of acceptance much more justifiable than the inflation version; such an actor has less of an autonomy interest, insofar as he has genuinely revealed an intention to commit a criminal act with the accompanying belief that harm was highly likely. (Notice the analogy to an actor who commits a criminal attempt.)

In one respect, Sleaze is not a good illustration of the contrast between the deflation and inflation scenarios, because once Sleaze hires the minor employee, his knowing decision to keep her on the payroll might well be a continuing crime. So whether he begins with the belief that she is very unlikely underage (under the deflation scenario) or with the belief that she probably is not (under inflation), at some point he does actually continue to have her in his employ when the state of his belief is that she is very likely under age. A closer analogy to Slumlord occurs if she is in his employ only for a short period of time, and, in the inflation scenario, he never forms the belief that she is probably underage until after she has left the job.

<sup>122</sup> See TAN 70-76= supra.

“relate back” to the pre-act decision period. Instead, the actor, prior to the moment of choice, reveals a willingness to act in the face of a very high risk.<sup>123</sup>

It is a close question whether, on balance, it is justifiable to employ the acceptance approach in this “modest inflation” scenario. For this scenario is still more troubling than the deflation scenario, in which the actor actually makes a choice to create the risk when he believed a harm (or prohibited circumstance) was likely. On the other hand, it is less troubling than the original inflation scenario; there, post-act statements or conduct provide a much less reliable basis for determining what the actor would have done, and a much less fair basis for interfering with the actor’s right to choose what to do.<sup>124</sup>

Factor (4) emphasizes that only a very short period of time elapses between (a) the moment when the actor reveals a willingness to take the action with a belief that the risk is very likely, and (b) the moment when he takes the action with the (merely reckless) belief that the risk is less likely.<sup>125</sup> Why is this important? Consider a variation of the original Slumlord case. First, as before, Slumlord takes the penultimate step, setting up the incendiary device, believing that the adjoining building is likely to burn. He has yet to take the final step, of pressing the remote control. As before, he receives a phone call revealing that the risk is not likely but only substantial. However, in this variation, he waits another six months before returning to the scene and pressing the button. Here, we should respect the actor’s autonomy and presume that he might well have had a change of heart if at that much later date he came to realize that the harm was, not merely substantial, but likely. And once again, the reason is based not only on the lesser reliability of any prediction of what he would have done, but also on a nonempirical value judgment: he is entitled to a presumption that, at a much later date, he might have acted differently.<sup>126</sup>

---

<sup>123</sup> The ambiguous facts of Michaels’ Sleaze example are open to a similar interpretation: perhaps, when Sleaze expressed his willingness to hire Minor even if she was definitely underage, he had not yet hired her.

To be sure, as an evidentiary matter, even an actor’s post-act statement that he would have been willing to act in the face of a high risk is somewhat probative of whether, prior to his act, he was, in fact, willing to act despite a high risk. Still, in many cases, post-act gloating or bragging will be insufficient evidence of pre-act acceptance. And, if one considers the pure inflation approach clearly objectionable, one way to ensure a firm boundary between the pure and modest inflation scenarios is to reject use of post-act statements and insist that the actor must have expressed his willingness to take the higher risk prior to engaging in the criminal act.

<sup>124</sup> An advocate of the acceptance approach might find this qualification too narrow. Perhaps the approach is defensible in all inflation cases (as well as deflation cases), and my concerns about unreliable counterfactual judgments are adequately addressed by taking seriously the government’s duty to prove each element of the offense (including the mental state of acceptance) beyond a reasonable doubt. See Michaels, *supra* note 5=, at =. However, accurate prediction and estimation is not the only problem here; an unqualified inflation approach also unjustifiably infringes the actor’s autonomy. See TAN 121= *supra*.

<sup>125</sup> In inflationary cases, moment (b) precedes moment (a).

<sup>126</sup> Then why not take the autonomy principle even more seriously, and always presume that an actor might have chosen differently at the very moment of choice? This approach has its adherents, but protecting autonomy to this extent would require enormous revision of our current criminal law. In the law of attempts, only complete attempts (in which the actor believed she had completed the conduct necessary to bring about the crime) could be punished. See Alexander & Kessler, *supra* note 112=, at 1168-1174. But the logic of this approach goes even further. Any crimes premised on future dangerousness (in the sense of the actor’s intention to engage in a future dangerous or harmful act)

The last two factors mentioned above are more rarely invoked, but they are still necessary conditions. Factor (5) requires that the change in the actor's subjective belief about the magnitude of the risk does not result from the actor's own conduct. The point is to exclude the following type of case. Mario aims in the direction of the victim in such a way that the risk of death is 90%; then he adjusts his aim in such a way that he decreases the risk to 2%. (He first decides to fire a bullet very close to Victor's head, to graze him; but then thinks better of it, and decides to shoot in the air just to scare Victor; but the ricochet still kills him.) Clearly it is not proper to punish Mario for a knowing killing; he was once willing to create a high risk of death, but now he is not willing to do so. On its face, the act of aiming the gun away from the victim strongly suggests that he is willing to impose no more than a small risk of death.

The last factor, (6), requires that an appropriate causal connection underlie the prediction that the reckless actor would have acted the same way even if he had known that the risk was likely. If the predictive test is to be a reliable criterion of culpability, it must demonstrate the actor's culpability in a causally apt way. For one can imagine unusual sequences of events in which the counterfactual predictive criterion is satisfied, yet the apparently "accepting" actor is not equivalent in culpability to a knowing actor.

For example, suppose that Rick is the partner of Sleaze and ordinarily does the hiring. The following is possible: (a) Rick hired Minor believing the risk was only one in twenty that she was underage, and (b) Rick also would have hired her if he believed the risk was very high (say, 80%), thus (c) Rick literally satisfies the "acceptance" or cognitive criterion. And yet the following sort of reason for (b) would undermine the rationale for using the criterion. Suppose that if (b) were true, i.e., Rick had believed that the chance of Minor being under age was 80%, then Rick would not have wanted to take this more serious risk of violating the law and would not have hired Minor, and would have so informed his partner Sleaze. But Sleaze would then have responded by threatening Rick with physical harm unless he hired Minor, and Rick would have succumbed to the threat and hired her. Whether or not such a threat would suffice to provide him a duress defense, Rick does not display a culpability equivalent to the original Sleaze, who willingly hired Minor without regard to the probability that she was under age.

Articulating precisely when the causal connection is appropriate or inappropriate is not easy, and not a task that I will attempt here. It is worth noting, however, that this difficulty also arises with other mental state connection requirements. One can invent odd counterexamples like the above for more conventional counterfactual tests too, including purpose and the doctrine that intoxication constructively supplies the mental state of recklessness (so long as a sober person would have been aware of the risk).<sup>127</sup> The prospect of such counterexamples is not sufficient to delegitimize the use of those other mental

---

would also be illegitimate, including such crimes as possession, burglary, kidnapping, and many theft crimes. See *id.* at 1193.

<sup>127</sup> For example, it could be the case that Irma, because of her voluntary intoxication, is unaware of a risk that a sober person would have noticed; yet the causal nexus is sufficiently unusual that it makes little to apply the usual constructive recklessness approach. (Suppose Irma's drunken condition prompts a bystander to play a joke on her, blindfolding her against her will; she then stumbles against the victim and is charged with reckless battery.)

states or doctrines, and similarly should not suffice to condemn the use of the counterfactual cognitive test.

Consider, now, a possible extension of Michaels' acceptance approach to those who are culpably ignorant, rather than reckless. Specifically, some courts and commentators (including this author) have supported extending that approach to a subcategory of culpably unaware (or "cognitively negligent") actors: such actors might be punished the same as cognitively reckless actors if they (a) are negligently unaware of a circumstance or result, but (b) would have taken the risk even if they were reckless, i.e., even if they were aware of a substantial risk that the circumstance existed or that the result will occur.<sup>128</sup> This approach differs from the usual modern hierarchy (exemplified by the Model Penal Code) which distinguishes, for purposes of grading, simply between those who are negligently unaware of a risk and those who (recklessly) are aware and take the risk. Under this possible extension of the acceptance approach, we would punish a subcategory of the negligently unaware the same as the reckless. (Similarly, Michaels' actual acceptance approach punishes a subcategory of the reckless—those who "accept" a circumstance or result—the same as knowing actors.)

What are the advantages and disadvantages of this extension of the acceptance approach?<sup>129</sup> Consider a further variation of the Sleaze example:

I-2. Sleaze (Negligently unaware) ("deflationary" scenario)

Suppose, as before, that Sleaze initially is informed of a substantial chance that Minor is under age. (His assistant Bouncer informs him that two of four I.D.'s were forged; Bouncer does not know which ones.) But, in this variation, Bouncer later tells Sleaze that he (Bouncer) has misremembered; none of the I.D.'s was forged. Sleaze doesn't check any further even though he easily could, because he doesn't care whether she is under age. He hires Minor.

Even if a jury believes that Sleaze was not reckless when he hired Minor, because he did not then believe that there was a substantial risk that Minor was under age, arguably he should be punished as seriously as someone who was reckless, i.e., who was aware of such a risk. Thus, if recklessness is the minimum mens rea for the crime of hiring an under age person, Sleaze arguably

---

<sup>128</sup> See Michaels, *supra* note 5=, at 969 n. 51 (collecting sources); Horder, *supra* note 5=, at 502-503.

Under Michaels' analysis, the acceptance test affects the grading of an offense (and indeed, only the grading differential between recklessness and knowledge). He does not endorse applying a counterfactual cognitive test to elevate culpable negligence to knowledge. Michaels, *supra* note 5=, at 962, n. 26. Presumably he would also object to applying a counterfactual cognitive test to elevate culpable negligence to recklessness.

<sup>129</sup> The doctrine that intoxicated defendants are constructively deemed reckless, see TAN 79-81= *supra*, sometimes illustrates reflects this expanded acceptance approach: for some actors who are unaware of a risk because of intoxication would, if aware, still have confronted the risk. Of course, this is not always the case. Still, intoxication is often an especially culpable cause of unawareness, and for that reason might justify treating the actor as severely as a reckless actor. See TAN = *infra* [next section].

satisfies that standard.<sup>130</sup> Some of the introductory examples could be analyzed similarly.<sup>131</sup>

However, this extension of the acceptance approach is more problematic than Michaels' original approach. Such an extension equates the culpability of a counterfactual mental state with that of an actual mental state even though the latter does not reflect a conscious choice.<sup>132</sup> At the very least, this extension should be limited to the "deflationary" scenario illustrated in the last Sleaze example. Thus, compare this "inflationary" scenario:

I-3. Sleaze (Negligently unaware) ("inflationary" scenario)

Sleaze originally hires Minor relying on information that all the I.D.'s show the applicants to be above age. However, he negligently relies on his untrustworthy aide Bouncer and fails to inquire further. Then, after employing her for a period of time, and after she has left the job, he is informed that actually, there was a 2 in 4 chance that she was under age. He indicates that he would have hired her anyway.

It would be far more problematic to permit criminal liability in this variation than in the prior one.

Although extending the cognitive counterfactual test this far is problematic, it is still possible that culpable ignorance can be a defensible form

---

<sup>130</sup> To be most persuasive, this extension of the acceptance approach should incorporate the same six criteria discussed earlier. If R is a reckless actor who takes the risk and S is a culpably ignorant actor (or one who culpably believes the risk is less than substantial), then the factors are as follows:

1. The only difference between R and S is that S is (culpably) unaware of the risk while R is aware.

2. S would take precisely the same action, if S had been aware. (Note: the prediction will often be more difficult here than in the case of acceptance.)

3. Deflation is a more justifiable scenario than inflation for employing a counterfactual prediction. (Realistically, it is unlikely that an actor's belief would "deflate" completely and be replaced by ignorance; more plausibly, a belief that a risk is substantial might deflate to a belief that it is insubstantial.) Moreover, under inflation, the concern about autonomy is arguably even greater, because the actor was not, at any time, prepared to commit the crime with at least some awareness of risk.

4. Deflation or inflation occurs over a very short period of time.

5. The change from a reckless belief (that the risk is substantial) to a less-than-reckless belief (either a belief that the risk is insubstantial, or a state of ignorance) does not reflect a change in the actor's risk-creating conduct. (And similarly for a change from a less-than-reckless to a reckless belief.)

Compare a case in which the deflation or inflation in belief does reflect such a change. For example, D aims his gun in a different direction, so that a risk of death that he initially believed was substantial he now believes to be insubstantial (though it is still unjustified and culpable). D should not be charged with the more serious initial risk that he knowingly created but then decided to reduce. (Or suppose that Sleaze checked additional forms and therefore no longer believed there was a substantial risk that the girl was underage.)

6. Again, we must exclude odd causal connections.

<sup>131</sup> For example, in Example E (The Indifferent Speeding Driver), suppose the driver sees the pedestrian at the final light. Compare a variation (E3) in which the driver is unaware that a pedestrian is in the zone of danger at the final light, but in which the driver would still have taken the risk (as he did when he ran the prior lights even though he then had noticed the endangered pedestrians).

<sup>132</sup> See Michaels, *supra* note 5, at 962 n. 26.

of culpable indifference under a different test. We will see in the next section that other culpable indifference tests are plausible, ones that don't rely solely on cognitive counterfactual criteria. (For example, if the very reason that Sleaze is ignorant of Minor's age is due to his own instruction that Bouncer never check I.D.'s, the culpability of giving that instruction might warrant treating Sleaze as culpably indifferent.<sup>133</sup>) So even if a straightforward extension of the acceptance approach to culpable ignorance and recklessness is too problematic, other persuasive reasons might exist for sometimes punishing the culpable ignorance form of indifference more than (non-indifferent) negligence, and perhaps as much as recklessness.

This analysis of the extended acceptance approach also helps explain why Ferzan's "Jane" is not a convincing counterexample to the acceptance approach. I repeat the example here:

J. Jane

[I]magine Jane who is driving her car, and fails to notice John in the crosswalk because she is too self-absorbed. Now, Simons would say that if Jane would have continued on anyway, had she in fact seen John, she should be punished as culpably indifferent, rather than simply negligent. But let us assume a little history between Jane and John. Jane, in fact, hates John. Indeed, Jane was driving to John's house, where she planned to take out her shotgun and murder John. Now, if we are playing the "what if" game, why stop at culpable indifference? Isn't it in fact more likely that Jane would aim at John and purposefully (not culpably indifferently) run him over?<sup>134</sup>

Concerning Ferzan's first point, that I would support punishing Jane at least for culpable indifference, I now have more hesitation, for the reasons noted above.<sup>135</sup> If Jane, just prior to the collision, had been aware that she was posing just as great a risk to a different pedestrian yet chose to take the risk, then she would exemplify the deflationary scenario, and an extension of the acceptance approach could be justifiable. (Recall the indifferent speeding driver example (E) from the introduction.) But in Ferzan's actual ("inflationary") example, Jane did not reveal acceptance of the risk at any time prior to the collision. Accordingly, as I have explained, it is especially problematic to employ the acceptance approach here.

Concerning Ferzan's second point, that in principle the counterfactual approach could justify convicting Jane of intentional murder, this reductio argument fails. As explained above, six conditions are necessary to the proper application of the acceptance test. The most salient here are:

- (1) The only difference between the actor's actual mental state at the time of the crime and his counterfactual mental state is a difference in the magnitude of the risk that the actor subjectively perceives.

---

<sup>133</sup> Notice that in Sleaze, Slumlord, and Jane (discussed immediately below), the actor need not have had a culpable reason for ignoring or underestimating the risk; rather, the actor's culpability consisted in his willing to take the risk had he been aware of it, or had he not underestimated it.

<sup>134</sup> Ferzan, *supra* note 9, at 622-623.

<sup>135</sup> Ferzan is correct, however, that my prior writings support her interpretation.

(2) We can confidently predict that the actor would take precisely the same action, whether he had the actual mental state (reckless) or the counterfactual one (knowledge) [or intention, in this example].

The first condition is clearly not satisfied. And one reason for this condition is to exclude application of a counterfactual acceptance test to intentions, because intentions and purposes affect culpability quite differently than cognitive states do.<sup>136</sup>

The second condition also might not be satisfied (even as modified). Although, by hypothesis, Jane predictably would have intentionally killed John, this apparently would have been a different act from the act she actually committed—accidentally running over a pedestrian who turned out to be John. For example, if she had seen him in the crosswalk and chose to execute her intention to kill him, presumably she would aim the car towards him in order to assure her success.

To be sure, it is not an easy matter to individuate acts.<sup>137</sup> And, to view the example in a light most sympathetic to Ferzan's point, suppose we can plausibly describe Jane's failure to brake at a particular time and location as the act or omission that causes John's death, either through Jane's culpable ignorance or as an execution of her intention to kill. (For the omission to be an execution of her lethal intention, suppose she sees John in the crosswalk immediately ahead but deliberately chooses not to brake.) However, even if we could indeed conclude that the "same" act (or omission) has occurred without regard to whether Jane committed it unintentionally or intentionally, this is mere coincidence. Jane's intention to kill is legally irrelevant for the sound reasons embodied in the traditional concurrence requirement, discussed above.<sup>138</sup> Her intention might even be a factual cause of his death, but neither this fact, nor the fact that she harbored an intention to kill while she was driving, means that any killing she brings about while driving is an execution of that intention. Motivational concurrence between the intention and the killing is also required, as we have seen, in order to assure a justifiable connection between that mental state and the relevant act or harm.<sup>139</sup>

The last topic in this section follows naturally from the previous topic: suppose the actor was not only unaware of a risk of harm, but was nonculpably unaware; and yet, if he had been aware of the risk, he would have taken it. Should the acceptance approach be extended to such an actor, permitting conviction?

---

<sup>136</sup> For further explanation, see TAN 116-118= supra.

<sup>137</sup> See generally Michael Moore, Act and Crime (1993).

<sup>138</sup> See TAN 29, 52-55= supra.

<sup>139</sup> It is possible, however, that Jane has committed attempted murder by the time that she is driving to John's house with the intention of killing him. See TAN 112= supra. But this reinforces the point that intentional murder liability is inappropriate: such an attempt should receive a lesser punishment, since Jane has not yet committed what she believes to be the last act necessary to bring about his death.

However, Ferzan also rejects liability for such incomplete attempts. See note 111=, supra. Ferzan's Jane example illustrates her broader objection to premising criminal liability on predictions of future conduct, both in attempt liability and beyond. See note 123= supra.

It should not. Expanding acceptance this far stretches the counterfactual approach beyond the breaking point. Consider yet another (!) Sleaze variation:

I-4. Sleaze (Reasonably or nonculpably unaware)

Sleaze is informed of a substantial chance that Minor is under age. (His assistant Bouncer informs him that two of four I.D.'s were suspiciously altered; Bouncer does not know which ones.) Sleaze nevertheless authorizes Minor's hire. But Sleaze's reliable assistant Clyde interjects and tells Sleaze that Bouncer is just pulling Sleaze's leg; none of the I.D.'s was altered. Moreover, Clyde has retained the I.D.'s overnight, and shows them to Sleaze, who reasonably concludes that they are not altered. However, Sleaze also states that he doesn't care whether she is under age. He then hires Minor who, as it turns out, is underage.

The basic problem with imposing counterfactual liability here, or in the even more troubling "inflationary" example in the footnote,<sup>140</sup> is that an especially firm line should be drawn between noncriminal and criminal conduct. The law should not punish an actor who has only shown a counterfactual willingness to act outside the law, but whose actual mental state is noncriminal. (Another problem is the difficulty of satisfying the second of the six conditions I identified as prerequisites to the acceptance test: it is normally especially difficult to predict what the actor would have done if he had been aware of a risk, in cases where, at the time of the action, the actor was not only unaware, but reasonably unaware, of that risk.<sup>141</sup>)

To conclude this section's discussion of cognitive counterfactual tests, I return to Alexander's ingenious Deborah hypothetical, noted above.<sup>142</sup> The example persuasively suggests that an extremely broad counterfactual culpable indifference test is objectionable. But the features of the example that make it compelling do not undermine the use of the cognitive criterion in "acceptance" cases, or even in some "culpable ignorance" cases.

Let us consider the hypothetical more closely. As stated, the hypothetical is ambiguous in some respects. For clarity, assume that Deborah is charged with a crime of "reckless operation of a motor vehicle," or "recklessly

---

<sup>140</sup> I-4 is a "deflationary" scenario. The corresponding "inflationary" scenario is, as usual, even more troubling:

I-5 Sleaze (Reasonably or nonculpably unaware) (inflationary)

Sleaze originally hires Minor, reasonably relying on information that all the I.D.'s show the applicants to be above age. (The information is provided by his reliable aide Rufus, who always carefully checks for alterations or forgeries). Then, after Minor has been employed for a period of time and has left the job, Rufus sheepishly informs Sleaze that actually, there was a 2 in 4 chance that she was under age (because two of the I.D.'s had clearly altered dates of birth, as in the original Sleaze example). Sleaze indicates that he would have hired her anyway.

In this scenario, again, we have a special concern about autonomy, magnified even further by the fact that Sleaze was not at any time prior to the action even culpably unaware of the risk.

<sup>141</sup> The last Sleaze variation assumes this problem away.

<sup>142</sup> Text at note 12= supra.

exceeding the speed limit.” And assume that the actus reus of the crime is satisfied, i.e., Deborah actually is speeding, for this makes the example parallel to the other examples we have been considering. (However, I will drop the latter assumption below.)

One important reason for the persuasive force of the hypothetical is this: Deborah apparently<sup>143</sup> is not culpable in believing that she is creating a low level of risk; thus, under the facts as she believes them to be, the risk would be justifiable and she would be committing no crime. Therefore, the hypothetical (even assuming she is speeding) resembles the last (“inflationary”) Sleaze variation noted above (I-4), in which Sleaze reasonably believes that Minor is above age, and hires her; and later, only after she has left the job, indicates that he would have hired her even if she was aware that she might be below age.<sup>144</sup> Indeed, even if we invoke a “deflationary” variation in order to make Deborah a more plausible case for liability, criminal liability would be unwarranted.<sup>145</sup> In all of these cases—Deborah, a deflationary variation of Deborah, and the two Sleaze variations (I-4 and I-5)—it is unjustifiable to convert a noncriminal mens rea into a criminal one. And this is unjustifiable notwithstanding a prediction that the actor would, counterfactually, still have taken the risk if she had been aware of it (or even if she believed it was likely).<sup>146</sup>

---

<sup>143</sup> I say “apparently” because, as stated, the example leaves open the possibility that Deborah might be culpable in believing that she is posing only a very low level of risk. Perhaps she does not look at the speedometer at all, and relies only on the speed of the surrounding traffic. But Alexander does not believe in criminal liability for negligence. See note 10= supra. This probably explains why his example mentions only to Deborah’s subjective perceptions of risk. Those who do not share his categorical rejection of criminal liability for negligence could support punishment of Deborah if they consider her habit of never looking at her speedometer sufficiently culpable.

<sup>144</sup> See variation I-5, in note 140= supra.

<sup>145</sup> Consider this variation:

G-2. Deborah “deflation” variation

Deborah initially believes that she is speeding. (Her speedometer suddenly becomes stuck. She comments to a passenger that she believes, based on the surrounding traffic, that she is speeding. Her passenger asks her not to speed, but she continues to drive at the same rate.)

Then, Deborah reasonably believes that she is not speeding. (Her speedometer becomes unstuck, and indicates that she is traveling well under the speed limit.)

The police pull her over for speeding. (The speedometer, although unstuck, turns out to be inaccurate, and she is indeed exceeding the speed limit.)

Since recklessness is required for speeding, Deborah lacks the mens rea for the crime. And it is also unjustifiable to convict her based on her counterfactual mens rea. To be sure, like Sleaze in variation I-4, she was, at a prior point in time, willing to act in a manner that would violate the law. But at the time that she was stopped, she did not believe she was speeding, nor did she believe there was a substantial risk that she was speeding.

<sup>146</sup> One asymmetry between this Deborah “deflation” scenario and the last Sleaze variation is that Deborah actually did commit the crime of speeding earlier, when her state of mind was reckless; while Sleaze did not commit the crime of hiring an underage person when his state of mind was reckless. But if the state only prosecutes Deborah for the acts that they observed when they stopped her, this is irrelevant. And if it is thought important whether she actually did commit a crime earlier, even one for which she is not being prosecuted, then suppose a further variation, in which she was not

Alexander's original example also derives some of its force from two other features. First, unlike the last variation, his example involves inflation, rather than deflation, of the actor's mental state. Since Deborah never, while she was actually driving, believed that she was (or might be) exceeding the speed limit, it is especially troubling to employ a counterfactual predictive test, for reasons already discussed.<sup>147</sup>

Second, in the original example, the actus reus is not satisfied, i.e., Deborah is not actually speeding. Whatever concerns one might have about punishing an actor with an unconventional counterfactual mental states, those concerns properly magnify if the actor also has failed to satisfy the actus reus. Indeed, attempt doctrine often requires a higher mens rea for an attempt than for a corresponding completed crime.<sup>148</sup> For a crime of recklessness, such as Deborah's reckless speeding or reckless operation of a motor vehicle, purpose or belief that one is speeding might ordinarily be required for an attempt.<sup>149</sup> The Deborah example is a little less troubling if she is actually speeding, though still, criminal liability would be unjustifiable, for the reasons already noted.

From the analysis in this section, two broad conclusions can be drawn. First, counterfactual criteria of mental states need to be defined quite carefully so that criminal punishment does not depend upon overly speculative inquiries and does not unjustifiably limit or burden a person's freedom to change his mind or to entertain a belief or other mental state that he does not express in action. Second, the acceptance approach should be limited in two ways—first, to cases in which the actor, at some point prior to his criminal action, either possessed the more culpable belief, or revealed that, had he possessed such a belief, he still

---

actually speeding when her state of mind was reckless (because, say, the speed limit on that stretch of the road was much higher than she believed it to be).

<sup>147</sup> It is possible, however, that Deborah exemplifies the less troublesome "moderate inflation" scenario. This would be so if, while driving, Deborah expressed a willingness to speed. She would exemplify a "pure" inflation scenario if she expressed that willingness only after she had completed her drive. See TAN 121-124= supra.

<sup>148</sup> For attempt, a heightened mens rea of purpose (or at least belief to a practical certainty) is ordinarily required with respect to a result element even if the completed crime requires only recklessness or negligence. The criteria governing circumstance elements are less clear. See Dressler, supra note 21=, at 385-386, 388-389, 408.

What if Michaels' Sleaze and Slumlord examples are converted into attempts? (Suppose Sleaze attempts to hire Minor but is apprehended before he does, or Slumlord's incendiary device malfunctions.) As attempts, the two examples do not implicate this problem of heightened mens rea, so long as belief that a circumstance or result is likely suffices for attempt liability. For Michaels points out that the "accepting" actor's culpability is equivalent to that of a knowing actor, and that equivalence justifies applying the acceptance approach to attempts as well as completed crimes. See Michaels, supra note 5=, at 1031-1033. However, if one further extends the acceptance principle, treating culpable ignorance as recklessness, then the heightened mens rea for attempts can make a difference. If even an actual reckless mens rea is insufficient for attempt liability, then culpable ignorance that is counterfactually equivalent to such recklessness will also be insufficient.

<sup>149</sup> One complication here is whether, in the crimes "reckless speeding" and "reckless operation of a motor vehicle," the mens rea of recklessness attaches to a result, conduct, or circumstance element. The statement in the text is most clearly correct if "speeding" and "operation of a motor vehicle" are characterized as result elements. I think that characterization is plausible, insofar as the actor brings about the excessive speed of his vehicle or the creation of objectively unjustifiable risks, but the matter is not free from doubt. See Robinson, supra note 52=, at 149-153.

would have run the risk (thus excluding what I have termed “pure inflation” scenarios); and second, to cases in which the actor was at least culpably ignorant of the relevant risk.

C. The idealized criterion: further analysis

Does the idealized criterion solve the problems with cognitive counterfactual tests? This section examines two different versions of that criterion—a single, general criterion, and an approach that embraces more specific multiple criteria of culpable indifference. Each version has distinctive advantages and disadvantages.

1. *A single idealized criterion of culpable indifference*

Section A above introduced the idea of an idealized counterfactual criterion, calibrated to the conduct of an idealized, non-indifferent person. Under such a criterion, an actor’s modest indifference is connected to his action just in case he would not have taken the action in question if he had shown socially appropriate concern for others. And an actor’s extreme indifference is connected to his action just in case the action he took was an extreme departure from the conduct we would expect of a non-indifferent person, demonstrating that the actor gave considerably less weight to the interests of the other than he should have. For example, the Russian roulette player (D) who chooses to endanger another’s life simply to obtain a personal thrill is extremely indifferent to the victim’s fate. And a statutory rapist (F) with no interest in whether the victim is below the age of consent displays more culpability than an actor (F2) who makes good faith inquiries but then unreasonably concludes that the victim is not below the age of consent.

This test has some possible advantages relative to the counterfactual cognitive test. First, for the reasons already discussed, it does not pose any “connection” problem; the only required connection is that the actor’s conduct be compared to, and fall short of, the idealized standard.<sup>150</sup> Second, the idealized criterion is broader than the counterfactual cognitive test; and this is an advantage if one judges the counterfactual cognitive test to be unduly narrow in some cases.<sup>151</sup> Third, the idealized test’s use of an explicitly normative criterion might be an attractive feature. Consider the views of the drafters of the Model Penal Code, endorsing a test of “extreme indifference to the value of human life” as the criterion for depraved heart murder:

[S]ome formula is needed to identify the case where recklessness may be found and where it should be assimilated to purpose or knowledge for purposes of grading. ... In a prosecution for murder ... the Code calls for the ... judgment whether the actor’s conscious disregard of the risk, under the circumstances, manifests extreme indifference to the value of human life. The significance of purpose or knowledge as a standard of liability is that ... purposeful or knowing

---

<sup>150</sup> See TAN 90= supra.

<sup>151</sup> For example, in the Russian Roulette example, the defendant probably would not have killed if he had known that death would likely occur (since the uncertainty of success is part of the thrill of the “game”), yet the actor’s culpability is closer to that of a murderer than of a person who commits reckless manslaughter. See TAN 96=.

homicide [ordinarily] demonstrates precisely such indifference... Whether recklessness is so extreme that it demonstrates similar indifference is not a question ... that can be further clarified. It must be left directly to the trier of fact under instructions which make it clear that recklessness that can fairly be assimilated to purpose or knowledge should be treated as murder...<sup>152</sup>

There is indeed some value in having a jury engage in a direct normative evaluation of the actor's conduct, and perhaps in premising that evaluation on a comparison to purposeful or knowing causation of harm. (This approach also raises serious vagueness problems, however, a point addressed below.)

Fourth, the broad normative evaluative judgment that this idealized test employs has a critical and distinctive feature: the judgment embraces the relative justifiability of the conduct in question. Specifically, under this test, the culpability of "indifference" can express the extent to which an actor has, or lacks, justification (or excuse) for his act. In this respect, the test is quite different from more conventional mental states.

For example, what distinguishes the Russian Roulette Player (D) who endangers a victim, from a speeding driver who knowingly creates a risk of death to a pedestrian, if the risk of death in the two cases is roughly the same? Indeed, murder liability might be appropriate for the Roulette Player even if he creates a risk of death no greater than the 1 in 38 probability of landing on a given number in the actual game of Roulette<sup>153</sup>; but murder liability would certainly not be appropriate if a speeding driver created only a 1 in 38 chance of killing a pedestrian. One essential difference is that the Russian Roulette player who selects an involuntary victim acts for a socially reprehensible end, to obtain pleasure from endangering another, while the speeding driver's end, to reach a destination more promptly, is not intrinsically objectionable. And one way of articulating this difference is in terms of the justifiability of the act: the Roulette Player acts for a completely unjustifiable reason, while the speeding driver acts for a reason that is sometimes justifiable (though under the circumstances, it is not).

But should the strength of the justification be part of the analysis of a mental state or culpability term such as "indifference"? After all, under the conventional modern model of mental states and culpability terms, the relative strength of a justification would not normally be considered relevant. Rather, the actor is either justified or not.<sup>154</sup> To be sure, the nature of the justification does differ, depending on which mental state is required. Thus, if an actor purposely or knowingly causes a social harm, his conduct is prima facie unjustifiable. More precisely, it is forbidden unless it fits within a recognized, separate defense,

---

<sup>152</sup> ALI, Model Penal Code & Commentaries, Part II, §210.2, comment 4, p. 21-22 (1980). The Code drafters observe: "[I]t seems undesirable to suggest a more specific formulation. ... [Unlike the Code's definition, various alternative formulations] "impair[]" "the primary purpose" [served by a definition] of communicating to jurors in ordinary language the task that is expected of them. The virtue of the Model Penal Code language is that it is a simpler and more direct method by which this function can be performed." Id. at 25-26.

<sup>153</sup> Suppose he blindly pulls one bullet out of a bag of thirty eight, containing thirty seven blanks and one live shell, and the bullet turns out to be the live one.

<sup>154</sup> Of course, at sentencing, the judge can consider the relative strength of a justification, insofar as she has discretion over the length of the sentence.

such as self-defense, law enforcement, or necessity. By contrast, if the actor recklessly or negligently causes a harm, lack of justification is relevant in a different way—it is part of the prima facie case. One is not prima facie liable simply for creating risks that one realizes are substantial, nor for creating risks that one should realize are substantial; the risks must also be unjustifiable.<sup>155</sup> Nevertheless, under this conventional model, justification is an all-or-nothing matter in either situation: either the purposeful or knowing causation of harm is justified (as a matter of defense) or it is not; either the creation of known risks (or risks that the actor should have appreciated) is justified (as part of the prima facie case) or it is not.<sup>156</sup>

However, this conventional model is too simple. Lack of justification is relevant to culpability in ways not captured by the model, because lack of justification is a question of degree. Some intentional murders not justified by self-defense or other explicit defenses are nevertheless less unjustifiable, and thus less culpable, than most other unjustifiable intentional murders. Mercy-killings are the best examples.<sup>157</sup> The doctrine of imperfect self-defense, for example, partially justifies an actor who honestly but unreasonably believes that he faces a deadly threat or that the threat is imminent: such an actor is liable for manslaughter, not murder.<sup>158</sup> And some reckless killings (for example, depraved heart killings) are more culpable than most other reckless killings, in part because the risky behavior is especially unjustifiable.<sup>159</sup>

The mental state category of culpable indifference can more flexibly consider the relative degree of justification in such cases.<sup>160</sup> In the case of

---

<sup>155</sup> Moreover, the interests that can render risk-creation justifiable encompass a wider range and can be lesser in weight than the narrowly defined interests (such as self-defense) that suffice to justify knowing or purposeful causation of harm.

<sup>156</sup> Culpability as to circumstances (rather than results) is analyzed similarly. Thus, if it is a crime knowingly to possess a gun, only limited defenses (such as necessity) are available. If it is a crime recklessly to possess a gun, then the state must show that the possession was unjustifiable. But in both cases, justification is an all-or-nothing inquiry.

<sup>157</sup> I assume for purposes of this discussion that the jurisdiction does not provide a statutory or constitutional exemption from murder liability for euthanasia.

<sup>158</sup> See Dressler, *supra* note 5, at 231-232.

<sup>159</sup> To be sure, a finding either of negligence or recklessness depends on the actor's conduct being unjustifiable. But once the actor's conduct is determined to be unjustifiable, neither culpability category can express the degree to which the conduct is unjustifiable. After all, recklessness differs from negligence only in the actor's awareness of the risk, not in the extent to which his action departs from reasonable conduct or is unjustified.

<sup>160</sup> This is not to say, however, that culpable indifference refers only to how unjustifiable the actor's conduct is. Compare Larry Alexander's view that the criminal law needs only a single mental state, "insufficient concern," reflecting both the actor's beliefs about risks and the actor's "desires" (in the sense of reasons for taking the risk). Alexander, *supra* note 9. Alexander helpfully emphasizes that the degree of justification is important in assessing an actor's culpability. However, his unitary view of mental states unduly flattens the multiple culpability criteria employed in the criminal law. See Joshua Dressler, *Does One Mens Rea Fit All?: Thoughts on Alexander's Unified Conception of Criminal Culpability*, 88 Calif. L. Rev. 955 (2000).

More specifically, Alexander's catch-all category of "desires" as "reasons for taking a risk" encompasses such disparate factors as intentions and purposes, on the one hand, and justifications, on the other. Moreover, within each category, a host of questions remain, questions that are concealed by Alexander's model. Thus, in analyzing intention and purpose, how should intention be defined? Does intention to harm always deserve a higher penalty than knowingly or recklessly causing harm? How do culpable desires and emotions (such as anger or racial bias) affect culpability, even though

murder, for example, “extreme indifference” (or the traditional language of “depraved heart”) can express the idea that a reckless killing that would otherwise be manslaughter but is especially unjustifiable deserves to be punished as murder. (An example is a death resulting from playing Russian Roulette.) And, in the case of manslaughter, “moderate indifference” can express the idea that a merely negligent killing that would otherwise not be manslaughter is elevated to manslaughter if, although the actor lacks (reckless) awareness of a risk, his reason for lacking such awareness is especially unjustifiable. (Possible examples include the statutory rapist who is unaware of the victim’s age because he makes no inquiries and does not care if she is underage; or a thief with a concealed, loaded gun who is unaware of the risk of death he is posing because he is preoccupied with completing the theft.) Finally, the absence of culpable indifference could even be employed as a doctrine of partial mitigation: one who kills another purposely or knowingly but for a somewhat justifiable motive might be viewed as lacking the “extreme indifference” to human life that is characteristic of most murders, and thus might be convicted of manslaughter, and not (as the conventional approach would provide) of murder. (Mercy-killing is a possible example.<sup>161</sup>)

However, an important objection to an “ideal person” criterion of culpable indifference is its vagueness.<sup>162</sup> The vagueness problem is indeed

---

they are not properly analyzed as intentions or purposes? In analyzing justifications, which interests does the law consider sufficient? What kinds of beliefs and motives must the actor possess with respect to the justificatory facts? Must these beliefs be objectively reasonable as well as subjectively held? Are justifications analyzed differently when they come in by way of defense rather than the prima facie case?

The concept of culpable indifference can take account of degrees of unjustifiability. But other factors can be relevant as well. Even an act that in some sense is clearly unjustifiable need not display culpable indifference. (Consider a driver’s momentary inattentiveness.)

<sup>161</sup> Mercy-killing is thus analogous to imperfect self-defense; under the latter doctrine, the actor’s subjective motivation of self-defense can lower the penalty even though the actor is not fully justified. See note 158= supra.

Similarly, culpable indifference can more flexibly accommodate mitigating circumstances or reasons that fall short of offering a full excuse. Suppose the actor commits a criminal act in response to another’s threat, but the threat is insufficient to warrant a full defense of duress. (Perhaps the threat is not of physical harm, or perhaps the actor subjectively responds out of fear but his response fails the objective component of the duress test.)

<sup>162</sup> Another possible objection is that this criterion isn’t really distinguishable from a test of negligence, at least if the latter test differentiates degrees of negligence (e.g., from simple to gross). On this view, culpable indifference is an otiose culpability term.

This objection seems misplaced, however. As conventionally understood, negligence differs from culpable indifference. The negligence standard ordinarily emphasizes an objective standard of care, and permits little inquiry into the actor’s personal capacity to meet the standard. It also normally encompasses such modest faults as inattentiveness or lack of skill. Some jurisdictions indeed consider ordinary tort negligence to be a sufficient level of culpability for some elements of some criminal offenses. Momentary inattentiveness while driving, and lack of skill in conducting an activity with resulting harm to another, both satisfy this minimal negligence standard, yet they need not amount to even modest “culpable indifference” to the interests of others.

For this very reason, however, culpable indifference (rather than tort negligence) is arguably the minimal culpability that criminal law should require. See Simons, Culpability and Retributive Theory, supra note 5=. An instructive comparison here is an analogous test of egalitarian duties. In equal protection doctrine, instead of asking merely whether government has purposely discriminated against a particular group, we might ask whether government was “indifferent” to their interests,

serious.<sup>163</sup> And this problem magnifies if the law employs an explicitly normative indifference test more widely than in the depraved heart murder context—for example, employing a “modest” (rather than “extreme”) indifference test for involuntary manslaughter,<sup>164</sup> or employing such a test for result or circumstance elements of other offenses, such as nonconsent in rape.<sup>165</sup>

Kim Ferzan articulates this concern in the following example:

K. David

[I]magine that David is an emergency room doctor who believes that it is critical for him to be at work on time. Let us now imagine two individuals, Victor and Peter. Victor is a pedestrian who will be hit and killed by David if he runs the light. Peter is a pedestrian who has been hit by a reckless driver. Peter’s life, let us assume, will be saved if and only if David gets to work on time. Of course, David does not know about Peter (or Victor) at the time he chooses to run the light. He just believes that he can save lives if he gets to work on time. Is David culpably indifferent? Does he care less about death than he should? Does he care less about injured pedestrians than he should? Does he care less about Victor than he should? Indeed, isn’t it a far simpler question to look at David’s choice, not his attitude about that choice, to determine if he is a culpable actor?<sup>166</sup>

Ferzan’s example nicely illustrates the danger of interpreting “indifference” merely as an isolated, acontextual subjective attitude of “not caring” about causing harm. Rather, any defensible conception must specify the types of choices and actions that exemplify “indifference,” and must be fully responsive to the moral context. If David is willing to run over a pedestrian in

---

relative to its concern for the interests of more favored groups. And one way to make this last question tractable is to ask whether a government that equally valued the interests of all groups would have enacted the policy at issue. See Simons, *Rethinking Mental States*, supra note 5=, at 521. Similarly, one way to articulate the lack of concern of a culpably indifferent criminal actor is to ask how the actor would have treated the victim if the victim had been one of the actor’s friends or family. However, this approach is inadequate insofar as it condemns actors for quite modest forms of insensitivity. Moreover, private individuals (in contrast to governments) can permissibly show greater concern for some persons to whom they have special ties. Thus, the fact that an actor would drive more slowly if his child or friend were his front-seat passenger does not demonstrate that driving at a slightly higher speed with a hitchhiker as a passenger reveals culpable indifference to the welfare of the hitchhiker. Criminal law should (and normally does) address only very serious deficiencies in concern for the interests of others.

<sup>163</sup> The passage above from the Model Penal Code, supra at note 152=, drastically underestimates this difficulty. For further discussion of the vagueness problem, see Michaels, “Rationales” of Criminal Law Then and Now: For a Judgmental Descriptivism, 100 *Colum. L. Rev.* 54, 75-79 (2000) (“Judgmental Descriptivism”); Michaels, *Acceptance*, supra note 5=, at 980-981, 1012-1013; Simons, *Culpability and Retributive Theory*, supra note 5=, at =.

<sup>164</sup> Suppose that involuntary manslaughter were defined simply as “causing the death of another under circumstances manifesting indifference to the value of human life” (employing a weaker version of the Model Penal Code’s “extreme” indifference murder statute).

<sup>165</sup> Suppose that the mens rea for nonconsent in rape were defined simply as “engaging in sexual intercourse with another under circumstances manifesting indifference to whether the other consents.”

<sup>166</sup> Ferzan, supra note 9=, at 624 (footnote omitted).

order to increase by a tiny amount the probability that he will get to the hospital in time to save a patient who would otherwise die, then his values are terribly misguided—in the same way, but perhaps a little less dramatically, than in the case of a terrorist who believes that killing innocent civilians for a political cause (a cause which, if successful, he believes will alleviate enormous suffering) is a morally justifiable way of expressing respect for life. Thus, however David himself would characterize his own values and attitudes, his actions reveal that he places much too low a value on the life of pedestrians whom he might kill.<sup>167</sup>

Still, my mere assertion that David actually “places much too low a value on life” does not make it so. The vagueness problem remains serious. It is also probably inevitable, if the idealized culpable indifference standard is to serve the two distinctive functions that it is expected to perform: first, to articulate an explicitly normative, rather than descriptive, culpability criterion<sup>168</sup>; and second, to serve as a kind of residual culpability term (relative to the narrower culpability criteria of purpose, knowledge, and cognitive recklessness).<sup>169</sup>

How is the vagueness problem to be solved? One possibility is to include “lack of justification” more explicitly in the standard. But that will only get us so far. Thus, if “extreme indifference” murder is defined as “causing another’s death for an especially unjustifiable reason” or “without even a pretense of justification,”<sup>170</sup> the vagueness problem is mitigated, but only slightly.

A second possibility is to explicitly define culpable indifference in relation to other mental states that are more clearly defined. In other words, why not ask directly whether the actor’s culpability is equivalent to the culpability of a purposeful, knowing, or cognitively reckless actor? Then extreme indifference could be defined as that degree of culpability that is comparable to the culpability of a purposeful or knowing actor. (For example, extreme indifference murder is a killing that displays a culpability similar to that of a purposeful or knowing

---

<sup>167</sup> Ferzan’s example also derives some of its force from a very different principle: directly creating a risk of death to a pedestrian can be wrongful even if one reasonably believes that creating the risk is the only way to lower the risk of death to one’s patient by an equivalent (or even greater) amount. Elizabeth Anderson gives an analogous example illustrating the moral inadequacy of unqualified consequentialist balancing: “A doctor may not neglect the health of her patient, a corporate executive whose demise will cause his firm to cease neglecting its workers’ health.” Elizabeth Anderson, Value in Ethics and Economics 73 (1993). Both examples reveal an important nonconsequentialist dimension to the morality of risk-imposition. See Kenneth W. Simons, *Negligence*, 16 Soc. Phil. & Pol. 52, 83 (1999).

However, in David’s case, we need not rely on this principle. For David very likely has alternatives (leaving earlier for work!), and his ex ante belief that his speeding significantly decreases the risk of someone dying is unlikely to be reasonable.

<sup>168</sup> As to the first function, the culpable indifference test is similar to any negligence criterion. See Michaels, *Judgmental Descriptivism*, supra note 163, at 65 n. 63; Pillsbury, supra note 5, at 83-86 (offering a distinction between “analytic” and “allusive” criteria that closely resembles the descriptive/ normative distinction). Thus, a culpable indifference standard will inevitably be more vague than descriptive criteria.

<sup>169</sup> See TAN 86= supra.

<sup>170</sup> Compare the language in the Model Penal Code comments explaining that the question whether a risk is “substantial” for purposes of the recklessness standard depends in part on the degree of justification for the risk: “[L]ess substantial risks might suffice for liability if there is no pretense of any justification for running the risk.” Model Penal Code §2.02(c), cmt 3, n. 14 (1985).

killing.<sup>171</sup>) And modest indifference would mean the degree of culpability that is similar to the culpability of a cognitively reckless actor.

This approach has promise, but it is also imperfect. For it presents an apples and oranges problem: Jurors would not have a clear idea of what they were comparing. After all, what is a “typical” knowing or purposeful killing? A related point is that some purposeful or knowing killings are much less justifiable or more blameworthy than others; and the same diversity occurs within the category of reckless killings.

## 2. *More specific multiple criteria of culpable indifference*

The vagueness problems with a simple idealized criterion are considerable. But a solution is possible—namely, to develop a set of much more specific criteria for culpable indifference. Although this solution loses something by way of comprehensiveness, it gains in clarity. (This tradeoff is, of course, common in other legal contexts when rules replace standards.) Moreover, these more specific criteria also largely avoid the “connection” problem, as we will see.

The following enumeration of five possible criteria of culpable indifference is meant to be suggestive, not exhaustive. The basic point is that these more specific criteria can fulfill much of the promise of a general idealized criterion—its explicit normativity, its attention to degree of justification, and its breadth relative to the cognitive counterfactual test—while minimizing problems of vagueness and of insufficient connection to the defendant’s acts.

### 1. Intent to risk and other especially inculpatory motives

One who intends to expose another to risk, as a means or end, is especially blameworthy.<sup>172</sup> Thus, in the Russian Roulette player example, the defendant is especially culpable, even if he does not satisfy the cognitive counterfactual test, i.e., even if we are not confident that he would have taken the risk had he known that the risk was certain to cause harm. And he might be more culpable than an actor who merely knowingly (and unjustifiably) created the same degree of risk, but for a less reprehensive motive.<sup>173</sup> Other aggravating motives (for pecuniary gain, out of bias, or “heinous, atrocious, or cruel”) can similarly be understood as articulating very specific kinds of “indifference” to the welfare of others. Such motives often serve as criteria for first degree murder, or for the death penalty in the post-liability sentencing stage.<sup>174</sup>

---

<sup>171</sup> See TAN 152 = supra; Michaels, Judgmental Descriptivism, supra note 163=, at 86. Analogously, actors who are wilfully blind to a circumstance could be defined simply as those whose recklessness and attitude towards the harm (as expressed in their conduct) display a culpability similar to that of a knowing actor. The actual judicial doctrine of wilful blindness is, of course, more elaborate than this.

<sup>172</sup> See Alexander, supra note 9=, at 938-939; Duff, supra note 5=, at 175-179; Simons, Negligence, supra note 167=, at 62, 82-83.

<sup>173</sup> See discussion at notes 96-97= supra.

<sup>174</sup> For discussion, see Pillsbury, supra note 5=, at 109-110; Carol S. Steiker, Punishing Hateful Motives, 97 Mich. L. Rev. 1857, 1866-1867 (1999). The presence of an especially culpable motive is considered one way to demonstrate that a murder was “especially heinous, cruel, or depraved.” See Steiker, id. at 1867.

Many instances of this first category (e.g. “intent to endanger”) present few vagueness problems, though other instances (e.g., “especially heinous, cruel, or depraved”) remain exceedingly imprecise. However, the connection requirement for “intent to expose to risk,” “a motive of financial gain,” and any other motive for action is straightforward and relatively unproblematic, as we have seen.<sup>175</sup>

2. Anger, racism, and other culpable desires and emotions that causally contribute to action (yet are not motives or intentions)

As noted above, a culpable desire can have legal significance even if it is not motivational—in other words, even if the motivating reason for the actor’s conduct was neither to accomplish the end desired nor to express the desire.<sup>176</sup> If an actor suddenly strikes a person of a different race in response to a racist impulse, he can justifiably be punished for a bias crime (even if he did not intend to express an insult to the person or his racial group). Or if sudden anger causes a driver to become distracted from risks on the road, he can justifiably be punished for culpable indifference to those risks (even if he does not purposely lose his temper or make a deliberate choice to express his anger). Or, as Samuel Pillsbury explains, the source of an actor’s unthinking rage can be an aggravating factor for murder—thus, “the butchering of a child for reasons of sexual frustration and rage” is extremely blameworthy.<sup>177</sup>

The connection requirement for such desires or emotions is not quite as straightforward as for the first category, but it is still not especially problematic. In these cases, the actor is blameworthy for failure to control her emotions, or her response to those emotions, when she could and should have exercised such control. However, if the emotion is not a factual and proximate cause of the action—for instance, because its effect has completely dissipated by the time of the action, or because the emotion arises only subsequent to the action—then criminal punishment based on possession of that emotion is not warranted, because the connection requirement is not satisfied.<sup>178</sup>

3. Participation in an immoral or illegal activity

An actor’s choice to participate in an immoral or illegal activity can be relevant in at least two ways to whether his acts demonstrate culpable indifference. First, such a choice can aggravate the culpability of his mental state or act, relative to an actor who does not so participate. For example, suppose an actor carries a loaded gun in his pocket in the course of committing a theft on the subway and is aware that he thereby is creating a substantial risk of death. Compare another actor who foolishly, and illegally, carries a loaded gun in his pocket on the subway while traveling to his target shooting class, aware that he is

---

<sup>175</sup> TAN 37-39= supra.

<sup>176</sup> See TAN 56-63= supra.

<sup>177</sup> Pillsbury, supra note 5=, at 104, discussing *People v. Anderson*, 447 P.2d 942 (Cal. 1968). However, Pillsbury misleadingly characterizes such aggravating factors as rage and cruelty as especially aggravating motives. As I have explained, it is better to analyze these cases as involving causal but not “motivational” desires and emotions. TAN = supra. See also Simons, supra note 58=, at 418-420.

<sup>178</sup> See the earlier discussion, TAN 56= supra.

creating the same substantial risk of death. If a death accidentally occurs, the first actor can justifiably be treated as more blameworthy, and as acting with a more serious degree of culpable indifference, than the second.

Second, as noted earlier,<sup>179</sup> if the reason why an actor lacks awareness of a risk is culpable, we might be justified in treating the actor as if he was aware of the risk. Participation in an immoral or illegal activity is one such culpable reason. Thus, if participation in an illegal activity explains the actor's lack of awareness of a risk, we might treat the actor as reckless. Furthermore, if his participation in such an activity explains his failure to infer from the facts that the risk is likely (and thus explains his possessing a reckless rather than knowing state of mind), we might treat the actor as knowing. Such a constructive approach parallels the legal doctrine that treats intoxicated actors as reckless when, but for their intoxication, they would have been aware of the risk.<sup>180</sup> Indeed, it is more justifiable to employ such an approach here, since it is normally far more culpable to choose to engage in a crime than to choose to get drunk.

This third category is relatively clearly defined (although not if the category includes "immoral" as well as "illegal" activities). With respect to the connection requirement, the category is not especially problematic. Thus, when the category applies in the first manner noted above, the actor's choice to participate in the activity is actually constitutive of the relevant action and thus helps explain how culpable the action is. And when it applies in the second manner, that choice causally explains his ignorance of the risk.

Of course, the relevance and weight of this factor need careful analysis, in order to ensure that the punishment is proportional to the actor's deserts. It is one thing to give modest weight to the choice to engage in a felony, and quite another to convert a case that would not otherwise justify any significant punishment into murder. Debates about the retributive justice of felony murder liability are directly pertinent here.<sup>181</sup>

#### 4. Successive creation of multiple risks over a short period of time

If an actor's course of conduct reveals a series of decisions to impose risks on others, he displays a concrete type of indifference to human welfare. And unlike a single decision to impose risks, a decision whose culpability can be mitigated if the actor acted on an impulse that he later regrets, sequential decisions are much less likely to deserve mitigation. Introductory example E illustrates this factor, as do a number of depraved heart murder cases.<sup>182</sup>

---

<sup>179</sup> See TAN 74-75= supra.

<sup>180</sup> See TAN 79-81= supra.

<sup>181</sup> See Kevin Cole, Killings during Crime: Toward a Discriminating Theory of Strict Liability, 28 Am. Crim. L. rev. 73 (1991); Kenneth W. Simons, When is Strict Criminal Liability Just?, 87 J. Crim. L. Criminol. 1075, 1120-1125 (1997); Dressler, supra note 5=, at 515-519.

<sup>182</sup> See, e.g., *People v. Gomez*, 478 N.E.2d 759 (N.Y. 1985) (defendant killed two pedestrians; he sped through busy New York City streets, struck two cars, twice mounted the sidewalk, and never applied his brakes).

5. Mitigating factors, including efforts to avoid harm and relatively benign motives

This last category is a negative criterion of culpable indifference: the presence of certain mitigating factors suggests that the actor is not culpably indifferent.<sup>183</sup> Some courts employ such a criterion. Thus, in involuntary manslaughter cases, courts adopting a “modest indifference” approach sometimes consider whether the actor made efforts to minimize the harm.<sup>184</sup> Some depraved heart murder cases, similarly, consider such efforts.<sup>185</sup>

Precisely how should this negative criterion apply? An instructive example is a New York depraved heart murder case, *People v. Roe*.<sup>186</sup> In this case, a 15-year old defendant killed his 13-year old friend while playing a game of Russian roulette. After causing the death, the defendant was observed kneeling over the body, crying, and later pounded his fists against the wall in anguish, exclaiming that he could not believe that he had shot his best friend. The defendant also immediately called for an ambulance.

Should the defendant’s reaction to the shooting affect whether he has satisfied the statutory standard of “depraved indifference to human life”? No and yes. The fact that he felt remorse, because he cared for his friend and did not want him to die, should not be relevant, for this retrospective emotional response has no causal bearing on the act of inflicting the fatal wound.<sup>187</sup> On the other hand, the defendant’s decision to call an ambulance is relevant, for it demonstrates his effort to prevent the fatal result from occurring,<sup>188</sup> and therefore

---

<sup>183</sup> See, e.g., Barry Mitchell, Culpably Indifferent Murder, 25 *Anglo-Am. L. Rev.* 64, 73, 80 (1996).

<sup>184</sup> See, e.g., *Commonwealth v. Agnew*, 398 A.2d 209 (Pa. Super. 1979), reversing a conviction of involuntary manslaughter where defendant towed oversized farm equipment on a road at night and caused a deadly collision. The court emphasized the efforts that defendant made to minimize the risk: he employed warning signs, drove slowly, and pulled his tractor over to a complete stop when the other vehicle was about to pass. “This is not an indifference to possible consequences, but a conscientious attempt to reduce the risk of an accident.” *Id.* at 212.

<sup>185</sup> See LaFave, *supra* note 23=, at 669-670. LaFave notes a case in which a highly reckless driver came to the assistance of the pedestrian he struck, bringing the victim to the hospital; the court held that these acts “negative the idea of wickedness of disposition and hardness of heart” that depraved heart murder requires. *Commonwealth v. McLaughlin*, 142 A. 213, 215 (Pa. 1928).

<sup>186</sup> 542 N. E. 2d 610 (N.Y. 1989).

<sup>187</sup> See also Ferzan, *supra* note 9=, at 625 n. 93, properly rejecting the argument, which she attributes to the dissenting opinion in the case, that defendant’s love for his friend negates depraved indifference murder.

However, defendant’s remorse would have some evidentiary bearing if defendant also claimed that he misunderstood the mechanics of the gun and honestly believed, before he fired, that there was virtually no chance that a live bullet would discharge. See note 18= *supra*.

<sup>188</sup> The dissenting opinion relies on this effort, but also on the defendant’s remorse, to criticize the majority’s conclusion that the evidence supported a finding of depraved indifference murder. 542 N. E. 2d at 617-618. An important complication is that New York courts have interpreted the statutory depraved indifference standard as requiring a greater objective risk of harm, but not a more culpable mental state, than reckless manslaughter requires. See majority opinion, *id.* at 612. Even on this (dubious) interpretation, however, calling an ambulance has some chance of reducing the objective risk; accordingly, evidence of this effort should be considered relevant to whether defendant has satisfied the depraved indifference standard.

is sufficiently connected to the fatal result to have a just bearing on his punishment.<sup>189</sup>

Examples of more benign motives have been discussed above—the motive to alleviate pain and respect the victims wishes, in a mercy-killing; and the motive of self-defense in a case where the actor’s subjective beliefs satisfy the legal requirements but those beliefs are considered objectively unreasonable. Similarly, the actor might have a mitigating reason for creating a risk, or for failing to appreciate the risk he has created—for example, he is in a state of understandable shock or grief.

A final question about this multiple criteria approach is whether it is really necessary. Could all of these criteria simply be assimilated within the “degree of justification” analysis presented earlier? Perhaps they could. For example, the actor’s decision to participate in an illegal activity makes her creation of an accompanying risk of death especially unjustifiable. However, separating the criteria out provides much more clarity and certainty than simply instructing a fact-finder to determine whether the defended acted “for an especially unjustifiable reason.” As a matter of drafting, a statutory implementation of this approach could either articulate the relevant multiple criteria, or provide a more general standard (such as “especially unjustifiable”), with the expectation that the judiciary would articulate the specific criteria over time, as experience developed.

#### IV. Conclusion

In order to avoid punishing merely for character or for attitudes unrelated to an actor’s criminal conduct, the criminal law must, when it requires proof of a mental state, employ connection requirements that appropriately link an actor’s mental state to his conduct. Even for the conventional mental states of purpose, intention, desire, knowledge, and recklessness, these connection requirements are complex and subtle. It should be no surprise, then, that the connection requirements for the unconventional mental state of culpable indifference must be defined with special care. However, if an explicit criterion of culpable indifference is to be justifiably employed in the criminal law, the challenge must be faced. Accordingly, and with these problems in mind, this paper demonstrates some advantages and disadvantages of two different versions of the culpable indifference criterion—a cognitive counterfactual criterion, which asks what harm the actor was willing to cause had his beliefs at the time of action been different; and an idealized criterion, which compares the actor’s conduct to that of a non-indifferent actor. However, defining culpable indifference simply as not caring whether a harm or legally prohibited state of affairs occurs is clearly inadequate. Neither indifference to harm in this very minimal sense, nor even

---

<sup>189</sup> A person’s effort to save a victim he has already injured is, in other words, more akin to the effort by the driver in *Agnew* (note 184=, supra) to avoid a collision in the first place, than to a case of pure remorse. However, if defendant in *Roe* knew that the actor was already dead or had no chance of survival, then his effort to call an ambulance is simply a concrete manifestation of remorse or respect for the victim, and thus is not sufficiently connected to his actions to preclude a criminal conviction that would otherwise be proper. See also note 31= infra (if remorse leads an actor to take the concrete step of renouncing an attempt, then it is indeed sufficiently connected to his actions that it can properly be considered in assessing whether he should be punished for attempt).

desire to harm, is sufficiently connected to action to be a proper basis for criminal punishment.

The analysis in this article reveals that my earlier model of mental states was incomplete and in some respects incorrect. That model proposed two mental state hierarchies (belief-states and desire-states) in lieu of the single hierarchy endorsed by the prevalent Model Penal Code. The model would justify the law employing (a) a hierarchical ranking among belief-states (e.g., acting with knowledge is considered more culpable than acting with cognitive recklessness), and (b) a separate hierarchical ranking among desire-states (e.g., acting with a desire to cause harm is considered more culpable than acting with indifference). However, a single ranking among all of these mental states—for example, the Model Penal Code hierarchy placing purpose, knowledge, recklessness, and negligence in order of decreasing culpability—is not defensible.<sup>190</sup>

The Code approach is indeed deficient, and its single-hierarchy ordering of all mental states is sometimes unjustifiable and sometimes incomplete. Most relevant here, it cannot explain the role of culpable indifference; although it does recognize the “extreme indifference” subcategory of murder, this is an ad hoc modification of the basic hierarchy. However, my earlier dual-hierarchy approach is inadequate, because it does not properly analyze conative states (of desire, intention, and purpose).<sup>191</sup> Specifically, the model improperly suggests that conative and cognitive states should be treated in a strictly analogous manner. Just as we can justifiably punish an actor who knowingly causes a harm more harshly than one who recklessly does so, the analysis suggests, we can justifiably punish an actor who commits an act with the desire to cause a harm more harshly than an actor who commits that same act but is indifferent to causing such a harm.<sup>192</sup> Unfortunately, however, the analogy is imperfect. It is indeed justifiable to differentiate distinct levels of culpability according to the actor’s different beliefs, as he acts, concerning the harmful effects of his actions (again, compare the knowing and reckless actor). But it is not similarly justifiable to differentiate distinct levels of culpability according to the distinct types of “desires” the actor possesses, as he acts. After all, acting in the face of one’s desire to cause harm need not be culpable at all; for one’s act might not be motivated by, or might not express, such a desire. (In the terms employed in this paper, the desires might not be appropriately “connected” to conduct.) Rather, the critical forms of culpability under the conative hierarchy are usually intentions and purposes. Mere desires are legally relevant only in limited circumstances—for example, where a fit of anger or a racist impulse is the cause, although not the motivation, of conduct.<sup>193</sup>

Similarly, an actor’s mere attitude of “indifference” towards causing harm, in the sense of the absence of a desire either to cause the harm or to avoid it, is not, by itself, relevant to retributive blame. Just as desires are ordinarily relevant only if they are incorporated within the motivational structure of intentions and purposes, “indifference” in this thin sense is not relevant; rather, it

---

<sup>190</sup> Simons, Rethinking Mental States, supra note 5=.

<sup>191</sup> For criticisms along these lines, see Michaels, supra note 5= at 963 n. 29, =; Ferzan, supra note 9=, at 620-627; Ferzan, =contribution to this volume=.

A full revision of the analytic framework of “Rethinking Mental States,” including a full response to Kim Ferzan’s critique in this symposium, =cite=, must await another occasion.

<sup>192</sup> See, e.g., Simons, Rethinking Mental States, supra note 5=, at 477.

<sup>193</sup> See TAN 55-63 supra.

must be connected to action in an appropriate way.<sup>194</sup> The project of this article has been to articulate the requisite connection.

---

<sup>194</sup> To be sure, occasionally desires are adequately connected to action merely as causes, at least where the actor could and should have controlled the desire. (Recall the discussion of anger and racist impulses.) Is it justifiable to employ a thinner conception of culpable indifference in a similar manner? In other words, if an actor's lack of preference or insufficient concern for the interests of others is a "cause" of his conduct, does that satisfy the connection requirement? Despite my tentative endorsement of this approach in the past, *Simons, Culpability and Retributive Theory*, supra note 5=, at 392-393, I believe that it is insufficiently subtle to capture the nature of the culpable indifference concept, and also insufficiently demanding to assure an adequate connection between the actor's mental state and his conduct.