Social Theory and Practice
Vol. 48, No. 3 (July 2022): 1–329
DOI: 10.5840/soctheorpract2021413123

The Real-Life Issue of Prepunishment

Preston Greene

Abstract: When someone is prepunished, they are punished for a predicted crime they will or would commit. I argue that cases of prepunishment universally assumed to be merely hypothetical—including those in Philip K. Dick's "The Minority Report"—are equivalent to some instances of the real-life punishment of attempt offenses. This conclusion puts pressure in two directions. If prepunishment is morally impermissible, as philosophers argue, then this calls for amendments to criminal justice theory and practice. At the same time, if prepunishment is not imaginary, then the philosophers who reject it cannot claim that their view is supported by common sense.

Keywords: Prepunishment, punishment, attempt, criminal law, compatibilism

In 1992, Christopher New published a provocative paper titled "Time and Punishment." He began by introducing a thought experiment in which the traffic police know that a person will exceed the speed limit tomorrow and decide to write a traffic ticket for the offense today. New claimed, i) that most people will judge this to be morally wrong, and ii) that this intuition is "mere prejudice." New granted that prepunishment may be imaginary, but he insisted that even if no one will ever be in a position to prepunish, "it is coherent to suppose that someone might" (36–37).

New's claim garnered a large number of responses. While most authors contributing to the debate disagreed that prepunishment is morally permissible, all accepted the assumption that prepunishment is imaginary. Daniel Statman (1997: 133) writes, "The possibility of prepunishment . . . is a purely theoretical one which, even if accepted, would have no implications for our human institution of punishment." Indeed, discussions of prepunishment in the philosophical literature add unrealistic assumptions, as in Smilansky 2007: 347: "Let us assume for the sake of our argument both determinism and complete predictability: if people's actions are determined, and we have perfect epistemic capacities, we can know ahead who will commit a crime." The only potential real-world application of prepunishment, according to Lloyd Strickland (2011: 108), is a divine one: "Needless to say, it is highly unlikely that

humans will ever develop sufficient foresight to make prepunishment a genuine option, with prescience seemingly forever to remain the preserve of God."[1]

That prepunishment is both weird and objectionable is a persistent thread that weaves its way throughout the prepunishment literature. As Stephen Kearns (2008: 253) writes, "Prepunishment is so bizarre that it can be resisted by just about anybody."

In this way, the philosophical literature on prepunishment mirrors the fictional literature. Many philosophers notice the similarity between New's case and Philip K. Dick's short story "The Minority Report," which is set in a fantastical future where prepunishments are doled out on the predictions of a trio of "precog mutants." In the film adaptation, "precrime" agents tap into the minds of these mutants to see visual representations of their predictions. This information is then relayed to John Anderton (played by Tom Cruise), who uses futuristic gadgets to apprehend the perpetrator before the crime occurs.

Saul Smilansky (2007: 347) explains that prepunishment occurs when a person is punished for a predicted crime at a point in time at which the crime has not occurred, whether or not the crime actually does occur (see also Beebee 2008: 259 and Wringe 2012: 135n4).[2] Accordingly, prepunishment can be either *preventative* or *non-preventative*. When it is non-preventative, as in New's example, a person is punished for a predicted crime and then allowed to actually commit the crime. When prepunishment is preventative, as in "The Minority Report," a person is punished for a predicted crime and not allowed to commit it.[3]

1. For other discussions of the significance of divine prepunishment, see Statman 1997: 133–34 and Todd 2013.
2. Smilansky continues:

> If the person does go on to commit the crime, and the only way of punishing him is through prepunishment, then prepunishment is the only way of establishing desert and justice. If prepunishment prevents the crime, it is morally tempting in a different way, because—unlike regular punishment, i.e., postpunishment—it is not inflicted after there are victims of crime, but rather prevents the crime, and so prevents also the potential harm. (2007: 347)

I accept Smilansky's conceptualization of prepunishment with one amendment. Sorenson (2006: 172) points out—persuasively in my opinion—that in using the term 'prepunishment,' we need not refer to the timing of the *punishment* but instead the timing of the *verdict*. In other words, what makes a punishment an instance of prepunishment is if the *decision to punish*, and not necessarily the punishment itself, occurs before the crime takes place.

3. In this article, I focus on preventative prepunishment because real-world prepunishment practices favor prevention when possible. However, an anonymous reviewer points out that in claiming that prepunishment is imaginary and morally impermissible, some philosophers have had in mind only non-preventative prepunishment. Nevertheless, the arguments made in this article remain relevant to these philosophers for the following reasons. First, philosophers arguing that prepunishment is currently impossible (see above for citations) have done so on the basis that we cannot have sufficient foreknowledge—in fact, that we are *nowhere close* to having sufficient foreknowledge—that someone will commit a crime. If this is true of non-preventative prepunishment, then it is also true of preventative prepunishment (because it is hard to believe that we could gain the sufficient foreknowledge simply by preventing what we predict). Second, the objections to prepunishment made by philosophers—even those philosophers who have only non-preventative prepunishment in mind—apply to preventative and non-preventative prepunishment equally (see section II and footnotes 24 and 26). Further, when they compare pre- and postpunishment, philosophers have universally concluded that preventative prepunishment is *at least as* morally objectionable as non-preventative prepunishment. Some authors claim that prepunishment is permissible when non-preventative and impermissible when preventative, but as far as I am aware, no one claims the reverse. For example, New (1993: 37) concludes that non-preventative prepunishment is permissible, but he grants that retributivists can reasonably reject the permissibility of preventative prepunishment. All told, then, our focus on preventative prepunishment serves to strengthen this article's conclusion that a practice that moral philosophers have claimed is morally impermissible but imaginary occurs in real-world legal systems.

Compared to related topics in moral and legal philosophy, prepunishment has so far enjoyed only modest popularity. If prepunishment were currently impossible, then this would, perhaps, be justified. However, prepunishment, or at least the part of it that matters morally, is not futuristic.

Here is a story of prepunishment that follows the formula in "Minority Report."

> *Science Fiction:* Precogs (mutants who reliably predict the future) report that John is plotting to murder his former wife, Joan, and will do so soon unless police intervene. Precrime agents arrest John at his apartment and take him to jail for the crime of "future murder."

This story seems fantastic and irrelevant to reality. But now consider the story of John as it might occur in the real world.

> *Real World:* John's friends report that John is plotting to murder his ex-wife, Joan, and will do so soon unless the police intervene. Detectives decide to follow John to try to determine whether John's friends are correct. They observe John as he drives to Joan's house. As he approaches the house, gun in hand, the detectives arrest John and take him to jail for the crime of attempted murder.

Real World is not fantastic; it has happened many times. What are the differences between the stories? Two differences stand out immediately. The first is that, in *Science Fiction*, evidence is provided by mutants rather than by John's friends and the observations of detectives. The second difference is that in *Science Fiction*, the detectives call the crime "future murder." Each difference is morally irrelevant.

First, it is simply a mistake to call the crime "future murder" because no murder ever takes place: John was about to murder his wife, but then the precrime agents stopped him. The thing the precrime agents care about in *Science Fiction* is better labeled "counterfactual murder." In this case, the precogs pre-

dict, counterfactually, what would happen if things were to play out without intervention (they are not "seeing the future" because what they see never occurs). In *Real World*, the detectives might make the same prediction: "what would happen if we do not intervene?" Accordingly, in each instance we might call the thing the authorities care about "counterfactual murder": if he had not been interrupted, then John would have murdered his wife.[4]

Second, while it is true that the *source* of the evidence differs between *Science Fiction* and *Real World*, it is only the *quality* of the evidence that matters. *Science Fiction* features evidence from mutants. *Real World* features evidence from John's friends and the observations of detectives. Mutants, testimony, and detective observations all supply evidence, in their respective worlds, for the same proposition: John will murder his wife unless the authorities intervene.

So far, we have shown that, at least in some instances, detectives might apprehend people for the same reason that precrime agents apprehend people in "Minority Report"—due to their beliefs about counterfactuals. So, insofar as the concept of "prepunishment" is expansive enough to include some of the things that might happen to a person before criminal conviction,[5] then prepunishment exists in the real world. However, this is not yet enough to show that punishments tied to criminal *convictions* are prepunishments. After all, in the real world, the crime is called "attempted murder" and not "counterfactual murder."[6] Is there a difference between the two?[7]

To answer this question, we will need to look at current legal theory concerning attempted murder. To anticipate: there is a view according to which the real-world prosecution of attempted murder should be analogous to that of counterfactual murder, and a view according to which they should not be analogous. To show that real-world prepunishment exists, we need not conclude that one view is "correct" and the other "incorrect." Instead, we will only need to show that each view is plausible and that each is sometimes consistent with the relevant statutes. If this is so, then we should expect that at least *in some instances*, the best interpretation of punishments for attempted murder is that they are punishments for counterfactual murder.

4. Or John would at least have assaulted his wife with the intent to kill. To be most precise, we could use the label "counterfactual murderous assault."
5. For example, see Wringe 2015 for the argument that "perp walks" qualify as a form of punishment.
6. There is an ambiguity in our use of the term "attempted murder." Sometimes it refers to situations in which someone assaults someone else, with the intent to kill, but the assault fails to kill. Other times "attempted murder" refers to situations in which someone is apprehended before an actual assault takes place. If attempted murder is sometimes treated in the same way as counterfactual murder, it is only in the latter sense.
7. I focus on attempted murder because this helps to narrow the potential considerations relevant to the discussion, while serving equally well to establish the conclusion that prepunishment sometimes occurs. Nevertheless, much of the discussion applies to other attempt offenses, as I sometimes note.

I

The statutes on attempted murder vary by jurisdiction, but many borrow heavily from the *Model Penal Code* (*MPC*).[8] Let's first look at the *MPC*'s guidance on attempted murder before considering some common alternatives. According to the *MPC*, attempted-murder charges require an intention to kill and that the offender took some "substantial step" toward the killing. Section 5.01(2) states that "substantial steps" may include any of the following (among others):

- "Enticing or seeking to entice the contemplated victim of the crime to go to the place contemplated for its commission."
- "Lying in wait, searching for or following the contemplated victim of the crime."
- "Possession of materials to be employed in the commission of the crime, which . . . can serve no lawful purpose of the actor under the circumstances."
- "Unlawful entry of a structure, vehicle or enclosure in which it is contemplated that the crime will be committed."

Consider a situation, like in *Real World* above, where a person intends to kill and takes a substantial step but is apprehended before the killing takes place. According to the *MPC*, such a person can be prosecuted for attempted murder. There are (at least) two interpretations of this: the *counterfactual* interpretation, which holds that the *MPC* aims to punish a person for what they would have done, and the *actual* interpretation, which holds that the *MPC* aims to punish a person for what they actually did. Let's consider the counterfactual interpretation first.

Supporters of the counterfactual interpretation would claim that in situations like *Real World* the *MPC* does not aim to punish a person for their actual intentions and steps, but for their counterfactual behavior: if the police had not intervened, the person would have committed a murder (or at least an assault with murderous intent). Intentions and steps provide evidence for the counterfactual: people often do what they intend to do, and they're more likely to do it if they've taken a substantial step toward it.

It is a familiar feature of legal systems that they must sometimes standardize what counts as sufficient evidence for certain kinds of claims for pragmatic

8. The *Model Penal Code* is a comprehensive synthesis of American law that individual states often consult in creating their penal codes.

reasons, even though such rules may not perfectly match an ideal epistemic account of evidential support in every case.[9] Helen Beebee (2008: 260), for example, writes, "The requirement that there be actual intentions and preparations can be seen merely as a reflection of the fact that these are the routes by which prosecutors come to have reasonable beliefs about what will happen if they do not intervene." In line with Beebee, a supporter of the counterfactual interpretation of the *MPC* would posit that its substantial-step test is a standardization of evidence for determining what a person would have done. After all, the *MPC* labels attempted murder an "inchoate crime" and explains that the punishment of such crimes *anticipates a further criminal act*. Showing that someone intended to commit a criminal act and took a substantial step toward its completion establishes that such an anticipation is justified.

What happens when an agent intends to murder and takes a substantial step but then changes their mind? If the *MPC* aims to punish for the relevant counterfactual and not the intention and step themselves, then we would expect a person who abandons their plan to be considered innocent. And this is exactly what we find. *MPC* Section 5.01(4) states that an agent cannot be convicted of attempted murder if he "abandoned his effort to commit the crime or otherwise prevented its commission, under circumstances manifesting a complete and voluntary renunciation of his criminal purpose." The counterfactual interpretation offers a natural explanation for this: the *MPC* does not aim to punish for the intentions and steps in themselves, since these are merely pieces of evidence for the relevant counterfactual. Abandonment is thus standardized as undermining the evidence supplied by the intention and step, and sufficient to establish that the person *would not have done it*. (Though as we'll see below, the actual interpretation has a competing explanation).

MPC 5.01(4) states that the renunciation of criminal purpose relevant to the abandonment defense cannot be due to the person encountering unforeseen circumstances that made the crime difficult to accomplish (or difficult to accomplish without being detected). This shows that the counterfactual interpretation of the *MPC* is a bit more permissive than stated above: a person is punished for attempt because they would have committed the crime had they not been stopped by the police *or* had they not encountered unforeseen circumstances that made the crime difficult to accomplish.[10][11]

More support for the counterfactual interpretation can be found in the work of Gideon Yaffe (2010). Yaffe's position on the prosecution of attempt diverges from the *MPC* in rejecting the abandonment defense (he claims that abandonment should instead be only a mitigating factor at sentencing). However, Yaffe focuses explicitly on counterfactuals in describing the nature of attempt. He proposes that an attempt is complete when it is the case that a person would commit the crime if they had the means and opportunity to do so. Yaffe's position thus agrees with the counterfactual interpretation of the *MPC*, but it broadens the relevant "completion counterfactual." According to Yaffe, an attempt is complete when an agent would commit the crime if they encounter no unforeseen circumstances, are not apprehended by the police, and do not abandon their intention.

Overall, the counterfactual interpretation makes a claim that many find plausible: when we punish for attempted murder, we do not punish a person for taking a step toward a crime while having a certain mental attitude. According to supporters of the counterfactual interpretation, to focus only on what the person has actually done would not fully capture the intuitive reasons the perpetrator is culpable for a heinous crime (since, at least according to the *MPC*, the substantial step might not even come close to harming anyone). To fully capture these reasons, we must appeal to counterfactuals; i.e., we must appeal

9. Such pragmatic considerations apply to many facets of the law beyond standards of evidence. As one example, David Brink (2017: 186) argues that the "bivalence" assumption concerning attempts, which holds that whether an attempt has been completed is an all-or-nothing question, can only be justified "for pragmatic reasons having to do with the limited ability of courts to track subtle differences in wrongdoing or culpability. If so, then bivalence might be defensible, if at all, as part of non-ideal theory, rather than ideal theory." (Nevertheless, there may be important methodological lessons that moral philosophers can learn from legal theorists. See Husak 2004).

10. Wringe (2012: 128) argues that if attempt law involved prepunishment, then we would never punish *bungled* attempts. But, pace Wringe, punishing bungled attempts is allowed under the "unforeseen circumstances" element of the *MPC*'s completion counterfactual (on the counterfactual interpretation). For example, a supporter of the counterfactual interpretation could claim that a bungled attempt should be punished because the perpetrator satisfied the relevant counterfactual: if they had not encountered circumstances that they did not foresee, they would have completed the criminal act.

11. One of the more vexing questions about attempts is whether the "impossibility" of an attempt matters. In some cases, it appears that impossibility shouldn't matter; for example, it doesn't seem to matter when a person intends to pickpocket but the pocket they choose contains nothing. In other cases, impossibility does seem to make a difference; for example, it seems to matter when a person intends to kill through "voodoo" and buys a doll for this purpose. Theorists disagree about how impossibility matters to attempt offenses, but the relevant point for our purposes is that the counterfactual interpretation seems best equipped to accept (or reject) versions of the impossibility defense through adjustments to the "unforeseen circumstances" clause. For example, money being absent from the pocket might count as an unforeseen circumstance that makes the crime more difficult to complete, while the impossibility of the doll causing harm to another person involves the basic causal structure of the world, which is a circumstance (that some might view as) relevantly different. The *MPC*, for its part, rejects all versions of the impossibility defense except for "legal impossibility," in which the attempted act, if consummated, would not be a crime. (See Duff 1997).

to the things a person would have done had they not been apprehended or had they not encountered unforeseen circumstances.[12]

Let's now turn to the actual interpretation. The actual interpretation rejects appeal to counterfactuals, and instead claims that a substantial step, together with the requisite intention, by themselves constitute the crime.[13] On the actual view, the crime being punished *just is* the criminal having the intention to kill while taking a substantial step. What the criminal would have done is irrelevant, and so there is no "completion counterfactual."

To explain the abandonment defense, a supporter of the actual interpretation might first note that criminal conviction establishes not just that a person acted wrongly, but that they acted wrongly and *it is appropriate to punish them*.[14] They might then claim that the abandonment defense shows only that punishment is inappropriate, and not that wrongdoing has not occurred.[15] (Indeed, on the actual view, an attempt is complete prior to abandonment if the person has taken a substantial step with the requisite intention).[16]

The *MPC* views abandonment as an affirmative defense, which means that it must be proven by the defendant (in this case, by merely a preponderance of

12. Considerations against moral luck may offer another line of support for the counterfactual interpretation. For example, Statman (1997: 130) argues that "whether our intention to act is realised or not depends on factors over which we have only limited control." If it is true, as he argues, that one's "real desert" is "immune to the contingencies of reality and which does not depend on what one actually does," then the counterfactual interpretation gives the best account of desert. (Statman (1997: 131) rejects the permissibility of prepunishment because of our "epistemic shortcomings," but supposed shortcomings aside, I take his argument to be conducive to the counterfactual interpretation, and thus to the permissibility of prepunishment.)
13. Wringe (2012: 128) labels what I call the "actual" interpretation the "distinct-offenses" interpretation. I prefer the actual/counterfactual terminology because it better highlights the crux of the question of whether attempt crimes involve prepunishment. Note that even supporters of the counterfactual interpretation can, and probably should, continue to view attempted murder and completed murder as distinct crimes. One involves punishment for a counterfactual, and the other involves punishment for a completed actual criminal act. These can be viewed as distinct crimes without problem, and the punishment for each crime might differ for pragmatic reasons.
14. See, for example, the work of John Gardner (2007).
15. For example, the "communicative expressivism" theory of punishment advocated by R. A. Duff (2001) argues that the point of punishment is to cause the offender to feel remorse for their wrongdoing. A defender of the actual interpretation could appeal to this theory in explaining the abandonment defense—if the defendant has already thought better of what they intended, then there is little need to cause them to feel remorse. For an analysis of prepunishment in light of communicative expressivism, see Wringe 2012.
16. Interestingly, even though he defends the actual interpretation, Brink (2017) rejects the strategy for defending it discussed here. He argues that it is implausible to claim that an attempt is complete before abandonment. Instead, he proposes that supporters of the actual interpretation should reject the "bivalence" assumption and accept that attempts can be partially complete.

the evidence). This fact can be explained by either the counterfactual or actual interpretation. Some affirmative defenses, like *insanity*, seem to amount to an "excuse" of wrongdoing; others, like *consent* or *self-defense*, amount to the claim that no wrongdoing has occurred.[17] Is the abandonment defense more like an excuse or a denial of wrongdoing? That is the crux of the disagreement between the actual and counterfactual interpretations when it comes to abandonment.[18]

Two further pieces of data may bear on the disagreement. The first concerns the severity of punishment for attempted murder: should the punishment be the same as for the completed crime? It seems like there is more justification for this on the counterfactual interpretation and less on the actual interpretation. In fact, the *MPC* 5.05(1) recommends that attempt be of the same grade and degree as the completed offense (with exceptions).[19] This recommendation offers some prima facie evidence in favor of the counterfactual interpretation; though, ultimately, *any* sentencing guidelines can be justified on either interpretation. Once we allow pragmatic considerations to be relevant to sentencing, there is nothing, in principle, that stops the supporter of the actual inter-

17. This, at least, accords with the plausible view that justified actions are not wrong actions. See, for example, Hurd 1999.
18. Affirmative defenses are sometimes partitioned into "excuses" and "justifications." However, as Brink (2017: 185–86) notes, legal theory texts—which are clear on the status of insanity, consent, and self-defense—leave it unclear how abandonment should be categorized. Perhaps, as the counterfactual interpretation seems to imply, this is because the justification/excuse distinction does not successfully partition the space of affirmative defenses. Indeed, as Simester and Sullivan (2000: 540) write, many legal theorists refuse to classify defenses using the justification/excuse partition because of "uncertainties about the definition and application" of these terms (quoted in Husak 2005: 557). The *MPC* 1.12(3)(c) agrees: it does not define affirmative defenses in terms of justifications and excuses, but instead as involving claims that are "peculiarly within the knowledge of the defendant." Following the *MPC*, the supporter of the counterfactual interpretation might claim that abandonment is an affirmative defense simply because it concerns claims for which the defendant is in the best position to supply evidence. For consider a legal system in which prosecutors must prove that a defendant did not abandon their plan. This seems less natural than the inverse because the defense is in a better position to supply evidence for abandonment than the prosecution is in a position to supply evidence against abandonment. (A similar explanation applies to the evolution of the insanity defense into an affirmative defense. Initially, the burden of persuasion was on prosecutors, but this changed when John Hinckley Jr. was found not guilty of the attempted assassination of Ronald Reagan in 1981 because the prosecution was unable to prove that he was *not* insane. After this, legislators switched the burden of persuasion to the defense).
19. It's worth noting that this feature is not unique to US law; many law systems recommend that attempts be punished nearly as severely as consummated offenses. For example, the UK standard reference text *Blackstone's Criminal Practice* states, "Even in cases where a low level of injury (or no injury) has been caused, an offence of attempted murder will be extremely serious" [495]. This fact counters an argument against the realism of prepunishment given by Smilansky (2008: 262–63). See footnote 20 for more.

pretation from arguing that murder and attempted murder should be punished similarly. And there is nothing that stops the supporter of the counterfactual interpretation from arguing that counterfactual murder should be punished less severely than actual murder.

The second piece of relevant data is that an attempt is never charged in addition to the completed crime. In contrast, solicitation and conspiracy, in most jurisdictions, can be prosecuted separately. This again offers some prima facie support for the counterfactual interpretation of attempt. After all, if attempt, like solicitation and conspiracy, could be prosecuted separately from the completed crime, that would be evidence for the actual interpretation. The fact that it cannot, therefore, indicates a counterfactual element to attempt that differentiates it from offenses like solicitation and conspiracy.[20] (Though, again, this is unlikely to settle the matter, since a supporter of the actual interpretation might argue that an attempted crime and the consummated offense should never be prosecuted together only because of pragmatic reasons that do not apply to solicitation or conspiracy).

Finally, it may be worthwhile to expand our view beyond the substantial-step test of the *MPC* and consider common alternatives. These are the *proximity* test (which analyzes the amount *left to be done* in the crime), the *res ipsa loquitur* test (which asks whether the defendant had *no other purpose* than to complete the crime the moment they stopped progressing toward its completion), and the *probable-desistance* test (which asks whether the defendant "crossed a line" past which it is probable that they would have completed the crime if there were no interruption).[21] The proximity test, like the substantial-step test, is ambiguous between the counterfactual and actual interpretations. However, the res ipsa loquitur test supports an actual interpretation, while the probable-desistance test strongly supports a counterfactual interpretation. The

20. A prior brief exchange between Beebee (2008: 259–60) and Smilansky (2008: 262–63) concerning the realism of prepunishment focused on UK conspiracy laws. Given that conspiracy can be punished in addition to the consummated offense but attempt cannot, a focus on attempt law provides a stronger starting point for establishing the realism of prepunishment. (For a view on the prepunishment status of UK conspiracy laws from a legal theorist's perspective, see Zedner 2007—cited in Tomlin 2015: 277). Furthermore, the fact that the *MPC* recommends that attempts be punished as severely as consummated offenses counters Smilansky's (2008: 263) appeal to sentencing laws. He claims that the difference between conspiracy and prepunishment is "illustrated through the very laws in question . . . in that the sentences for conspiracy and for its implementation will almost invariably differ." By Smilansky's own reasoning, then, the *MPC*'s recommendation that the punishment of an attempt crime should not differ from that of the consummated crime is at least prima facie evidence that the *MPC*'s doctrine on attempts does involve prepunishment.

21. For more on these tests and citations of relevant cases, see "Attempts" in *Criminal Law* (University of Minnesota Libraries Publishing, 2015). Locatable at: http://open.lib.umn.edu/criminallaw/.

probable-desistance test supports a counterfactual interpretation because it references counterfactuals directly: is it probable that the defendant would not have desisted had they not been stopped? Meanwhile, the res ipsa loquitur test seems to take a non-counterfactual approach: what was the defendant's actual state of mind when they stopped progressing toward the crime?

Generally speaking, whether we should punish people for their counterfactual behavior is a difficult theoretical dispute with prominent defenders on each side. I suggest that what makes the problem hard is conflicting intuitions: some people feel that mere intentions and steps in themselves fail to fully capture the nature of the crime, and their punishment may border on "thought crime" (i.e., "we're punishing you not just because you had an intention and took a step, but because you were *going to commit the crime*") and others are against the punishment of counterfactuals (i.e., "we're punishing you for what you *did*, and not for merely what you *would have done*").[22] These dueling motivations have resulted in statutes that are sometimes more in line with the counterfactual interpretation and other times more in line with the actual interpretation, and sometimes ambiguous between the two.

However, the key question for our purposes is not whether it is the counterfactual or actual interpretation that succeeds in describing the way things should be, but instead which interpretation best describes the way things are. The jurisdictional variations in the statutes, combined with the ability of each interpretation to coherently explain the motivations behind most statutes, strongly suggest that in real practice the prosecution of attempted murder is sometimes in line with the counterfactual interpretation and sometimes in line with the actual interpretation. To deny this, one would have to make the very strong claim that in *no* instances does the counterfactual interpretation best describe the successful prosecution of attempted murder. This is hard to believe, especially given those jurisdictions that use the probable-desistance test. (And we should not forget that in some states attempt is not criminalized in a statute, but is instead left to the variable results of common law).[23]

22. An excellent example of this dynamic can be seen in Alexander and Ferzan's *Crime and Culpability* (2009). They argue that if we take seriously the view that "an actor should be punished only for what he *has done* and not *what he will do*" (198), then we should never punish people for intentions or steps that do not, in themselves, risk harm to another. Though revisionary in restricting the criminalization of attempt, their theory does an excellent job of avoiding punishing for either thoughts or counterfactuals. The important point, for our purposes, is that Alexander and Ferzan claim that modern legal theory and practice concerning attempts *does* punish people for counterfactuals, and hence they call for revisions.

23. As an anonymous reviewer points out, it is possible to formulate a hybrid version of the actual interpretation. Specifically, this would be the claim that when a person is punished for attempted murder, they are punished for what they actually do (e.g., taking a step while having an intention), but whether it is appropriate to punish them depends on whether they

As we saw above, the instances of prepunishment described in "Minority Report" concern the punishment of counterfactual murder (or counterfactual murderous assault). Real-world attempted-murder punishments, according to the counterfactual interpretation, punish for the same crime. The only differences lie in the sources and standards of evidence used to establish the relevant counterfactual. Therefore, if the counterfactual interpretation best describes some instances of the successful real-world prosecution of attempted murder, then prepunishment exists in real life.

II

If the above is correct, then the significance of "Minority Report" has been underestimated. Dick is not merely presenting a story about a futuristic world; he's inviting us to see in a new light some of the workings of our punishment practices as they actually exist today. Dick imagines bizarre "precogs" that supply more reliable evidence than witnesses, detective observation, or substantial steps ever could, but he leaves untouched all the features of the situation that are morally relevant. And the interesting thing is that the view of prepunishment taken by many philosophers implies that what Dick describes is morally wrong.

As previously noted, New claimed that resistance to prepunishment is the result of "mere prejudice." Since New's article, however, a seeming consensus

would have murdered someone had they not been apprehended or encountered unforeseen circumstances. (As we saw above, the actual interpretation already relies on the claim that abandonment entails that a person acted wrongly but that it is not appropriate to punish them. The current suggestion would extend that explanation to the defendant's counterfactual behavior). On this hybrid interpretation, it is still the case that the person's actual behavior is being punished for, but their counterfactual behavior is a necessary condition for such punishment to be appropriate.

I believe this is a coherent interpretation worth taking seriously. At the same time, we should notice that the counterfactual interpretation admits of a corresponding hybrid formulation. On this formulation, when a person is punished for attempted murder, they are punished for what they would have done, but whether it is appropriate to punish them depends on whether they actually took a step and had an intention. On this hybrid counterfactual interpretation, it is still the case that the person's counterfactual behavior is being punished for, but their actual behavior is a necessary condition for punishment to be appropriate. I believe the hybrid counterfactual interpretation is at least as plausible as the hybrid actual interpretation. Further, we should note that once we move to these hybrid interpretations, it becomes increasingly difficult to separate their plausibility by looking at the attempted murder statutes. Each view implies the same rules regarding abandonment, significant steps and intentions, and counterfactuals. The fact that these interpretations are very difficult to separate by looking at the statutes underscores the claim that we should not assume that the real-world punishment of attempted murder is always best interpreted as the punishment of actual behavior, as opposed to the punishment of counterfactual behavior (and thus, by hypothesis, prepunishment).

has formed that prepunishment is immoral. In his response to New, Smilansky (1994: 51–52) writes:

> While in postpunishment the offender cannot take back her actions, in prepunishment she still has time to choose. She can decide, even in the last minute, not to commit the offence. . . . This explanation [for why prepunishment is wrong], it seems to me, is that in prepunishment we are not showing the respect due to the moral personality of the agent, who is, when 'punished', as yet innocent, and who we *must respect as capable of not committing the offence*. In prepunishment there is categorically still time, a 'window of moral opportunity' for the would-be offender.

Considering the preceding discussion, we can interpret Smilansky not as merely providing an "explanation" for something everyone already believes. Instead, he is arguing against the counterfactual interpretation of attempt offenses. According to Smilansky, we must refrain from punishing people on the basis of counterfactuals: doing so is unacceptable if a person had any time left to change their mind—even if we have very strong evidence that they were not going to. Therefore, if Smilansky is right, then the counterfactual interpretation of attempt offenses, and the real-world practices that conform to it, are misguided.

Other arguments against prepunishment focus on the relationship between free action and prediction. Fred Feldman (1995: 75) writes:

> Consider a typical case in which it seems quite likely that a certain person will commit some crime. We think he will deserve the legally mandated punishment only if he will be responsible for the crime and we think that he will be responsible for the crime only if he will commit it 'freely'; and we think that if he will commit it 'freely', then it cannot yet be quite certain that he will commit it. There must still be some possibility that he will decide not to commit it. So we insist upon a legal system that prohibits punishment-in-advance.[24]

As with Smilansky, Feldman would be mistaken to think that he is explaining something everyone already believes. In fact, the US statutes, at least in some jurisdictions, *do not* prohibit the kind of punishment he identifies (at least, not for murderous assaults we are nearly certain would have been freely done had they not been interrupted). If Feldman is right, then these statutes are misguided.[25]

As a third example, consider Roy Sorenson's (2006: 170) argument against prepunishment: "My thesis is that a crime justifies a verdict only by being a cause of that verdict. The asymmetry of 'A causes B' explains why we can-

24. Notice that Feldman's argument, like Smilansky's, applies equally against preventative and non-preventative prepunishment.

25. Interestingly, Feldman argues against the claim that people do not *deserve* punishment for their future actions. Feldman claims that legal punishment is different: even though people might deserve punishment for future actions, they should not be criminally punished for them (and, according to him, our legal systems do not punish in this way).

not prepunish wrongdoers. In practice, causal asymmetry matches temporal asymmetry. This accounts for the appearance that temporal prepunishment is unjust."[26] Again, this argument cannot be merely an account of existing beliefs, but instead an argument for a particular theory of how we should conceive of attempt offenses. If the unconsummated part of an inchoate crime cannot justify a verdict, as Sorenson claims, then counterfactuals about what would have occurred had the police not intervened cannot justify attempted-murder verdicts. But, if our discussion above is correct, then such counterfactuals are sometimes seen to justify attempted-murder verdicts.[27]

I won't take a firm stand on whether these arguments against prepunishment succeed. The key point, for our purposes, is that these arguments need not be seen as merely concerning a possibility. They can instead be located within an important and far-reaching dispute of real-life legal theory and practice. The philosophical arguments against prepunishment serve equally as arguments against the permissibility of the counterfactual interpretation of attempt offenses. If the philosophical arguments succeed, then we should work to ensure that our legal theory and practice is never in line with the counterfactual interpretation. This could involve the wholesale rejection of the probable-

26. Notice that Sorenson's argument applies equally against preventative and non-preventative prepunishment. If a predicted crime cannot cause a verdict when it is allowed to take place, then such a crime also cannot cause a verdict when it is not allowed to take place. Therefore, for the arguments presented in this essay, it does not matter whether Sorenson has non-preventative or preventative prepunishment in mind.

27. Actually, in light of the discussion above, I suggest that Sorenson need not reject prepunishment. If, as Sorenson assumes, prepunishment verdicts were justified by a 'future crime,' then this would indeed be unjust according to his causal theory of verdicts, since future events cannot cause past events. However, prepunishment verdicts need not be seen as justified by future events but instead by predictions. What causes these predictions are observations of the present and past *state of the world*. So, one might claim that it is the present and past state of the world that both causes and justifies prepunishment convictions. For preventative prepunishment convictions, the state of the world causes just the conviction, while for non-preventative prepunishment convictions, the state of the world is a common cause of both the conviction and the future crime.

Similarly, I suggest that Wringe (2012) need not reject prepunishment. Wringe argues that the point of punishment is to induce the offender to feel remorse for their wrongdoing (i.e., a "communicative expressivism" view of punishment), and in the case of prepunishment the offender has not yet done anything wrong. He concludes, "it is logically impossible to feel remorse for something that one has not done" (130). However, if, for example, it is true of someone that *they will murder their spouse* or *they would have murdered their spouse had they not been apprehended*, then that, it seems to me, is an appropriate thing for which a person might feel something like remorse (even if it is true that, strictly speaking, this would not fall under the concept of 'remorse'). Consider the shame and self-condemnation you might feel if such were true of you. If so, then a broad communicative expressivist view of punishment may have the resources to accept prepunishment. (A similar argument for the permissibility of prepunishment—though not specifically designed for communicative expressivism—is proposed by Statman (1997: 133).)

desistance test, as well as the amendment of many other statutes to make them unambiguously in line with the actual interpretation. For example, it should be clarified in the *MPC* that the abandonment defense provides an excuse for prior wrongdoing.

At the same time, if prepunishment is not imaginary, then that does seem to exert some pressure against the philosophers who claim it is morally wrong. To reject prepunishment, it is not enough to point to merely imagined circumstances, or a science-fictional fantasy, and claim that what happens there is immoral. One must point to the *counterfactual interpretation of attempt crimes*, and call that immoral. Though many reject the counterfactual interpretation, its rejection does not seem to be an established element of common sense, in the same way that critics of prepunishment have taken their view to be so established. As I have tried to show in this article, there is, at least, some plausible support for the counterfactual interpretation.[28]

III

Overall, acceptance of the counterfactual interpretation of attempt crimes, and thus prepunishment, would leave us with a tidy explanation of our real-world punishment practices. The explanation would be this: we sometimes find ourselves wanting to punish people who haven't yet harmed anyone. According to supporters of prepunishment, the best explanation for this lies not in our dislike of intentions and preparations, but in what we fear these people would have done had we not intervened. I find much to like in this outlook on prepunishment.

Compared to related topics in moral and legal philosophy, prepunishment has so far enjoyed only modest popularity. If prepunishment were a mere possibility, then this would be justified. In fact, this is false: prepunishment is a real feature of legal systems as they exist today. This calls for changes to our thinking about the philosophical importance and moral status of prepunishment. Theorizing about prepunishment is a critical topic of inquiry for moral philosophers and legal theorists.

If prepunishment is morally impermissible, as philosophers have argued, then modern legal systems should be reformed in the ways outlined above. At

28. It is worth noting that there exists a separate and rich line of inquiry concerning the relationship between the impermissibility of prepunishment and free will—most prominently, whether it is consistent with compatibilism that prepunishment is morally impermissible. This debate started with Smilansky (2007) and has also garnered a significant number of responses. See, for example, Robinson 2010, and for further references, Lemos 2012. While this debate is outside the scope of this article, to the extent that the arguments presented here support the view that prepunishment is morally permissible, they also support compatibilism.

the same time, the fact that prepunishment is an element of actual legal theory and practice should cause those opposed to prepunishment to re-examine their conviction that they have the common-sense view.

Nanyang Technological University, Singapore
pgreene@ntu.edu.sg

Acknowledgements

This paper is dedicated to the memory of John Williams, who introduced me to the philosophical debate over prepunishment. (See Williams 2012.) The Singapore Ministry of Education Academic Research Fund Tier 1 provided funding for this work.

References

Alexander, Larry, and Kimberly Kessler Ferzan, with Stephen J. Morse. 2009. *Crime and Culpability: A Theory of Criminal Law*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511804595

American Law Institute. 2002 (1981). *Model Penal Code*. New York: Foundation Press.

Beebee, Helen. 2008. "Smilansky's Alleged Refutation of Compatibilism," *Analysis* 68(3): 258–60. https://doi.org/10.1093/analys/68.3.258

Brink, David O. 2017. "The Path to Completion," in *Oxford Studies in Agency and Responsibility*, vol. 4, ed. David Shoemaker. Oxford: Oxford University Press

Duff, Antony. 1997. "Impossible Attempts," in *Criminal Attempts*. Oxford: Clarendon Press. https://doi.org/10.1093/acprof:oso/9780198262688.001.0001

Duff, Antony. 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press.

Feldman, Fred. 1995. "Desert: Reconsideration of Some Received Wisdom," *Mind* 104: 63–77. https://doi.org/10.1093/mind/104.413.63

Gardner, John. 2007. "In Defence of Defences," in *Offences and Defences: Selected Essays in the Philosophy of Criminal Law*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199239351.001.0001

Hooper, Anthony, and David Perry. 2013. *Blackstone's Criminal Practice*, 23rd ed. Oxford: Oxford University Press.

Hurd, Heidi M. 1999. "Justification and Excuse, Wrongdoing and Culpability," *Notre Dame Law Review* 74(5): 1551–74.

Husak, Douglas. 2004. "What Moral Philosophers Might Learn from Criminal Theorists," *Rutgers Law Journal* 36(1): 191–98.

Husak, Douglas. 2005. "On the Supposed Priority of Justification to Excuse," *Law and Philosophy* 24: 557–94. https://doi.org/10.1007/s10982-005-0845-8

Kearns, Stephen. 2008. "Compatibilism Can Resist Prepunishment: A Reply to Smilansky," *Analysis* 68(3): 250–53. https://doi.org/10.1093/analys/68.3.250

Lemos, John. 2012. "Libertarianism and the Wrongness of Prepunishment," *Public Affairs Quarterly* 26(3): 243–55.

New, Christopher. 1992. "Time and Punishment," *Analysis* 52(1): 35–40. https://doi.org/10.1093/analys/52.1.35

Robinson, Michael. 2010. "A Compatibilist-Friendly Rejection of Prepunishment," *Philosophia* 38: 589–94. https://doi.org/10.1007/s11406-009-9236-y

Simester, A. P., and G. R. Sullivan. 2000. *Criminal Law: Theory and Doctrine*. Oxford: Hart Publishing.

Smilansky, Saul. 1994. "The Time to Punish," *Analysis* 54(1): 50–53. https://doi.org/10.1093/analys/54.1.50

Smilansky, Saul. 2007. "Determinism and Prepunishment: The Radical Nature of Compatibilism," *Analysis* 67(4): 347–49. https://doi.org/10.1093/analys/67.4.347

Smilansky, Saul. 2008. "More Prepunishment for Compatibilists: A Reply to Beebee," *Analysis* 68(3): 260–63. https://doi.org/10.1093/analys/68.3.260

Sorensen, Roy. 2006. "Future Law: Prepunishment and the Causal Theory of Verdicts," *Nous* 40(1): 166–83. https://doi.org/10.1111/j.0029-4624.2006.00605.x

Statman, Daniel. 1997. "The Time to Punish and the Problem of Moral Luck," *Journal of Applied Philosophy* 14(2): 129–35. https://doi.org/10.1111/1468-5930.00049

Strickland, Lloyd. 2011. "God and Prepunishment," *Philosophical Papers* 40(1): 105–27. https://doi.org/10.1080/05568641.2011.560029

Todd, Patrick. 2013. "Prepunishment and Explanatory Dependence: A New Argument for Incompatibilism about Foreknowledge and Freedom," *The Philosophical Review* 122(4): 619–39. https://doi.org/10.1215/00318108-2315315

Tomlin, Patrick. 2015. "Should Retributivists Prefer Prepunishment?," *Social Theory and Practice* 41(2): 275–85. https://doi.org/10.5840/soctheorpract201541215

Williams, John N. 2012. "Beyond Minority Report: Pre-Crime, Pre-punishment and Pre-desert," *TRANS: Internet Journal for Cultural Sciences* 17: 10–15.

Wringe, Bill. 2012. "Pre-Punishment, Communicative Theories of Punishment, and Compatibilism," *Pacific Philosophical Quarterly* 93(2): 125–36. https://doi.org/10.1111/j.1468-0114.2012.01424.x

Wringe, Bill. 2015. "Perp Walks as Punishment," *Ethical Theory and Moral Practice* 18(3): 615–29. https://doi.org/10.1007/s10677-014-9545-5

Yaffe, Gideon. 2010. *Attempts*. Oxford: Clarendon Press. https://doi.org/10.1093/acprof:oso/9780199590667.001.0001

Zedner, Lucia. 2007. "Preventive Justice or Pre-Punishment? The Case of Control Orders," *Current Legal Problems* 60: 174–203. https://doi.org/10.1093/clp/60.1.174