The poster boy for my paper is the King’s Messenger in Lewis Carroll’s *Through the Looking Glass*. Recall that since the White Queen lives backwards, her memory works forwards. She pities Alice who can only remember things after they happen. Alice asks which things the Queen remembers best:

‘Oh, things that happened the week after next,’ the Queen replied in a careless tone. ‘For instance, . . . there’s the King’s Messenger. He’s in prison now, being punished: and the trial doesn’t even begin till next Wednesday: and of course the crime comes last of all.’ (Chapter V)

Is the King’s Messenger being treated unjustly? Why not prepunish people for their future crimes? True, it is generally easier to have knowledge about past crimes than future crimes. But is unpredictability the only problem with prepunishment?

My answer is a causal theory of verdicts that condemns all actual prepunishment but for reasons that are temporally neutral. These reasons echo the semantic concerns of logicians who study the liar paradox.

1. Prepunishment

In 1956 Philip K. Dick published “The Minority Report”, a science fiction story about “pre-crime” (which was turned into a movie by Stephen Spielberg in 2002). Police harness clairvoyants to predict murders and then arrest the “perpetuators” before they commit their crimes. When Precrime Commissioner, John T. Anderton, is himself accused of the future murder of a man unknown to him, he becomes disenchanted with the system. He struggles heroically to prove his innocence.
This science fiction tale is conceptually problematic in distractingly many ways. How can a clairvoyant see an event that has yet to occur? Since ‘see’ is a success word, one can only see what is actual—yet the “seen” felonies are prevented.

When Alice asks what happens if the King’s Messenger never commits the crime, the White Queen replies, “That would be all the better wouldn’t it?”

Alice felt there was no denying THAT. ‘Of course it would be all the better,’ she said:
‘but it wouldn’t be all the better his being punished.’
‘You’re wrong THERE, at any rate,’ said the Queen: ‘were YOU ever punished?’
‘Only for faults,’ said Alice.
‘And you were all the better for it, I know!’ the Queen said triumphantly.
‘Yes, but then I HAD done the things I was punished for,’ said Alice: ‘that makes all the difference.’
‘But if you HADN’T done them,’ the Queen said, ‘that would have been better still; better, and better, and better!’

If Alice had cared to continue the debate, she might have distinguished the issue of whether the punishment improved her character from the issue of whether the punishment was just. If the King’s Messenger is being punished for stealing tarts, then it is better (more just) if he actually stole them.

The problem about punishing the innocent is graver in “The Minority Report”. For the judges know that the crime is prevented. This complicates John Anderton’s quest to prove his innocence. Since one cannot be guilty of a near miss of a crime, what could count as proving one’s innocence of a hypothetical crime?

The characters in “The Minority Report” evince fitful awareness of these conundrums. They allude to the subtleties of self-defeating prophecies, the metaphysics of branching futures, and the tricky semantics of ‘innocent’.

Fortunately, Christopher New’s (1992) defense of “prepunishment” evades the clutter of Philip Dick’s science fiction. New gives the example of Algy, an eccentric Alaskan motorist who informs the Sheriff that he will break the speed limit on Wilderness One at 10:31 tomorrow morning. Algy offers to pay the fine now but warns that he will evade any fine that is imposed after the crime is committed. The Sheriff knows Algy will do exactly as he says. So the Sheriff agrees to prepunish Algy.

Bear in mind that the Sheriff is not penalizing Algy for intending to speed. We do punish some attempts—but only after the attempts take place. We do not punish people for their future attempts. True, mere preparation for certain kinds of crimes is itself a crime. But it is a different crime than the one prepared for. And police still wait for the preparation to occur before making an arrest.
Daniel Statman (1997) interprets Christopher New as expressing opposition to “moral luck”: Given that we know that Algy intends to do wrong and know that he will not alter that intention, we have all that we need to justify punishing Algy. Intentions are what matter, not whether they are successfully carried out. William of Ockham and Immanuel Kant devised a number of thought experiments to underscore their intention-based ethics. Under Statman’s interpretation, New’s prepunishment proposal should be situated within this genre.

I disagree. The issue of prepunishment can be raised with an agent who merely foresees his crime. When Ulysses has himself tied to the masthead, he does not intend to direct his crew toward the deadly Sirens. He foresees this breech of his duty as captain (just as a realistic dieter predicts breeches of his diet). Ulysses might request to be punished before he sets sail (in port where authorities can oversee his punishment). Even if one analyzes Ulysses’ weakness of will as an intentional action, the punishment of Ulysses would occur prior to his bad intention.

When envisaging Algy, Christopher New was merely trying to construct an epistemically plausible example of prepunishment. People are easier to predict when they sincerely report intentions. I interpret New as an advocate of temporal neutrality. Since New’s point is that the timing of punishment is irrelevant, he should accept prepunishment even prior to formation of an intention to do wrong.

If New were unwilling to adopt the neutrality I have attributed to him, a pre-Calvinist would make a handsome substitution. Regular Calvinists believe God foresees our deeds and pre-ordains whether we will end up in heaven or hell. As we live, we gain clues about our fate by how well we behave. Evidential decision theory (epitomized by the first edition of Richard Jeffrey’s *A Theory of Decision*) instructs us to maximize good news and so is committed to endorsing the Calvinist’s efforts to raise his subjective probability of being amongst the elect. Causal decision theorists condemn this as auto-propaganda. What matters for causal theorists is producing objective improvements, not cheering yourself up by moving into a nicer reference class. (That’s why causal decision theorists choose both boxes in Newcomb’s problem.)

What’s intriguing about Calvinists is that they cut the causal link between sin and divine punishment. So now it is only a distracting nicety to schedule hell after the sin. Let us imagine the theology of a “pre-Calvinist” who believes heaven and hell come before this life. Each of us has an infinite past but from birth, a limited future. Each of us has either emerged from heaven or emerged from hell but no one remembers which. We gain clues about our infinite biographies by how we fare in the finite remainder of our existence. Thus a life filled with good works is evidence that you have spent an infinite amount of time in heaven. Or so says the pre-Calvinist.
Most philosophers are compatibilists and so will go along with Christopher New’s assumption that Algy’s free choice is predictable. We take the predictability of crime for granted in parole hearings. It has even been suggested to me that the American mixture of heavy prison sentences and liberal parole is a system of prepunishment.

Many people balk at the notion that free choice is predictable. In the movie version of “Minority Report”, critics of the precrime system emphasize that their free choice keeps the future open. These critics confuse predictive determinism and metaphysical determinism. Fidelity to the moral law might make metaphysically undetermined agents predictable. Complexity might make metaphysically determined agents unpredictable.

Metaphysical indeterminism should be distinguished from Saul Smilansky’s emphasis on moral opportunity:

. . . 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 offense. In prepunishment there is categorically still time, a ‘window of moral opportunity’ for the would-be offender. This moral opportunity needs to be acknowledged. In postpunishment, by contrast, it is too late. We cannot—and thus cannot be under any obligation to—allow for the agent’s e.g. last-minute moral improvement. (1994, 52)

Presumably, even if God knew that the agent will commit the crime, He must not prepunish out of respect for the fact that the agent could refrain from the wrongdoing.

Christopher New complains that such forbearance “would be no more than an empty gesture.” (1995, 60) There is only a point in waiting for something to happen if we do not know the outcome. New also observes that this Kantian respect for moral personality does not extend to legal “persons” such as corporations. Why not prepunish them?

And what about respecting morally punctilious people who request prepunishment, say, for their upcoming suicide? Georg Hegel credits the guilty with a right to be punished. From the Hegelian perspective, refusal to prepunish is a wrong inflicted on the guilty.

Lastly, Christopher New notes that there is no disrespect in pre-rewarding. He thinks it would be natural for commanders to award kamikaze pilots medals for gallantry before they complete their missions. New’s point can be amplified by citing the old custom of forgiving your executioner. In 1587 Mary, Queen of Scots, did not slight her executioners when she granted their request to forgive them for what they were about to do.

A final point: The boldness of New’s defense will be obscured if we blur the distinction between punishment and other impositions such as
involuntary quarantine. In 1907 an Irish immigrant, Mary Mallon, was identified as a healthy carrier of typhoid fever. As a cook, she had infected 22 people with one fatality. The New York City Health department quarantined her on North Brother Island for three years. She was released after promising to never again work as a cook. “Typhoid Mary” was later traced to Sloan Maternity Hospital working as a cook (25 more people infected with 2 deaths). She was returned to North Brother Island and died of a stroke 23 years later in 1938.

Although the legal basis for Mary’s initial quarantine is provocatively murky, Professor New is only interested in punishment. He concedes that it is usually advantageous to punish crimes after their commission. But Professor New denies this should be a universal requirement. He believes some punishments may precede their respective crimes.

2. A causal theory of verdicts

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 cannot prepunish wrongdoers. In practice, causal asymmetry matches temporal asymmetry. This accounts for the appearance that temporal prepunishment is unjust.

A verdict is a finding of guilt. The defendant’s legal status only changes when the verdict is incorporated in a judgment of conviction. (Convicted felons cannot vote, hold office, serve on juries, etc.) After the verdict, the judge gives the defense some time for legal maneuvers such as a motion to throw out the verdict. If the defendant dies during this interlude, he has been found guilty without being convicted. The defendant would be eligible to be buried in a national cemetery.

The causal theory of verdicts does not require that the penalty be caused by the verdict. Consider defendants who are detained pending the completion of their trials. If the defendant is sentenced to imprisonment, the judge gives the defendant credit for the time served while awaiting trial. For instance, if the defendant was detained for a year and was sentenced to two years, then his year of detainment is counted as a year of punishment. Had he been acquitted, the year of detainment would not be punishment. An acquitted defendant can subsequently testify that he was not punished during the period of his detainment. The guilty individual who makes the same claim commits perjury.

Clearly (yet surprisingly) there is prepunishment in the sense of there being punishment that precedes the verdict. I take Christopher New to be defending prepunishment in the sense of there being punishment that precedes the crime.

The causal theory of verdicts behaves like the simple causal theory of knowledge. Consider the legal counterparts of Gettier counterexamples. Suppose a judge has a justified belief that Jones murdered Smith.
Although the judge has no reason to think that Jones murdered Brown, he infers ‘Either Jones murdered Smith or Jones murdered Brown’. As it turns out, Jones did not murder Smith but did, coincidentally, murder Brown. On the causal theory of verdicts, the judge did not justly rule that ‘Either Jones murdered Smith or Brown’ even though the judge did have a justified true belief in this disjunction.

Gettier cases are often discussed in conjunction with Gilbert Harman’s (1973, 143–44) social knowledge cases. Jill reads about an assassination of a civil rights leader in a reputable newspaper. She never hears a government spokesman later announce that the victim of the shooting was actually a bodyguard. Unbeknownst to the public, this announcement is a lie designed to avert a coup. Although Jill’s reasoning does not include the sort of false step common in Gettier scenarios, Jill fails to know that the civil rights leader was assassinated. Or so says Harman. The case is controversial. The causal theory of knowledge takes sides; it implies that Jill does know. After all, there is a normal causal chain leading from the assassination to her belief in the assassination.

I personally think the causal theory of knowledge is wrong to disagree with Harman. Jill does not know. But this does not damage the causal theory of verdicts. Suppose Jill is a member of a sequestered jury. The jurors’ access to government announcements is legally blocked, so the jurors do not hear about the spokesman’s denial. They find the assassin guilty. This verdict was caused by the assassination in the normal way. Although the jurors do not know that the assassin killed the civil rights leader (according to Harman, me, and others), their verdict is still just.

The most damaging objection to the causal theory of knowledge is Carl Ginet’s barn counterexample: A father drives through the countryside with his son. For the boy’s edification, the father points to a barn and says ‘That is a barn’. Unbeknownst to the father, he is driving through a region in which there are many fake barns. Although the father pointed at a real barn, the neighboring facsimiles would have deceived him. The causal theory of knowledge mistakenly implies that the father knows that the structure he pointed at was a barn. After all, the causal process was the normal one of observation and there was a normal causal chain running from the barn to the belief it is a barn.

The causal theory of verdicts takes the fake barn cases in stride. Suppose that the father is a juror in a trial and is unaware that psychologists are running fake trials in the neighborhood. The jurors and “jurors” believe that they are participating in real trials. Thus the father does not know that the defendant is guilty. But his guilty verdict is still just because it is caused by the crime in a normal way. Justice does not require knowledge.

The simple causal theory of knowledge mistakenly implies that we have no knowledge of the future or of general statements or mathematical statements. Alvin Goldman (1967) tries to repair these defects by replacing
straight causation with causal connection (and by restricting the theory to empirical knowledge). The causal theory of verdicts does not involve this sophistication. It welcomes the implication that there can be no just verdicts about future events (putting aside reverse causation) or abstract entities or generalities.

As with causal theories of perception and memory, the causal theory of verdicts states a necessary condition, not a sufficient condition. Judges and juries strive to make their verdicts appropriate effects of past crimes. Deviant causal chains do not provide a basis for just verdicts. Suppose Murray Door murders Vic Tim with a knife. He cleans the knife and removes it from the crime scene. The murder of Vic Tim causes the police to frame Murray Door; they move his knife to the crime scene and stain it with Vic Tim’s blood. Thus the crime causes the jury to form a justified true belief that Murray Door murdered Vic Tim and Murray Door receives a punishment that fits his crime. But his punishment is unjust because of the abnormal causal chain. The police cannot shore up procedural justice by “veridically framing” defendants.

To reiterate, my treatment of prepunishment is directed at verdicts rather than the timing of the sanction. In Norway, there aren’t enough prison cells for all the drunk drivers, so there is a six-month wait to get into prison. Suppose Judge Odin foresees that Captain Loki will drive home drunk in six months after he docks. In deference to the Captain’s shipping obligations, Odin pronounces Captain Loki guilty six months before he commits his crime so that Loki’s prison sentence will commence immediately after his crime. I count this as prepunishment because I am dating by verdicts rather than by the execution of the penalty.

Frankly, I regard the penalty phase as a dispensable distraction. There is no penalty for breaking some smoking bans. Even smokers who know that the law is “without teeth” refrain from breaking it. Consider a psychically gifted judge who foresees that Miss Puff will violate the smoking ban tomorrow but renders his verdict today. Every body knows that there is no penalty attached to the guilty verdict. In my book, this would be prepunishment even though there is no sanction.

In torts, the focus is compensation rather than punishment. Yet concerns about causal pedigree penetrate. Suppose I concede that you were justified in docking your boat to my dock to escape a sudden storm. I can still sue you for the damage your boat does to my dock while buffeted by the waves. And I would probably win given the precedent of Vincent v. Lake Erie Transportation Co. (10 Minn. 456, 124 NW [1910]). Further suppose that a judge forecasted all this and ruled in my favor before you docked your boat. If you object to prepunishment, you should object to precompensation. You will not be mollified if the judge schedules the compensation to be paid after your boat damages my dock.
The demand for causal evidence excludes reasoning from duplicates. In 1981, the huge ship Derbyshire sank with hardly a trace in a typhoon. Had the crew been negligent? The U.K. Department of Transport commissioned a study of its sister ships. All were discovered to have the same structural flaw, a construction error that was the basis of a promising hypothesis as to how the Derbyshire sunk. However, a formal inquiry in 1987 ruled this evidence inadmissible. Relatives of the dead crew were outraged. Whether or not the evidence was conclusive, they felt it should be weighed. To get admissible evidence, the seamen’s union was forced to finance an expedition to find the wreck of the Derbyshire.

Our sympathy with the crew and their relatives may lead us to prefer a legal reform in which the causal condition is relaxed. There is already some evidence of this relaxation in torts; statistical arguments have been admitted in assessing corporate liability (Thomson 1986). In criminal law, the insistence on causation has been unbending. Reasoning from duplicates is never allowed. If we discover that we have convicted the identical twin of the real perpetuator, we require a whole new trial for the new suspect. The prosecution cannot skip the trial by proving that the first trial would have had the same outcome if the correct twin had been at the defense table.

Notice that the causal theory of verdicts is compatible with probabilistic evidence. For much probabilistic evidence satisfies the causal requirement. Trouble comes from the species of statistical evidence that skips the crime and exploits precursors of the crime.

3. What’s wrong with prepunishment?

Suppose Judge Farsighted convicts a woman for an assault she will commit upon release from prison. Judge Farsighted foresees that the woman will bitterly resent her conviction and seek revenge against him. Upon release she indeed attacks Judge Farsighted. When the police arrest her, Judge Farsighted orders the woman released on the grounds that she has already served her sentence for the assault.

The problem is not that Judge Farsighted verdict causes a crime. Judges know many of their guilty verdicts cause crimes. Prison corrupts many bad people into becoming worse people. A consequentialist judge can still justify the punishment on the grounds that it will deter others. A retributive judge can still justify the punishment on the grounds that, regardless of consequences, the criminal deserves the punishment.

The problem with Judge Farsighted’s verdict is that it causes the very crime it purports to punish. There is no independent grounding for the punishment. Judge Farsighted lets the justification for the judgment loop back on itself. When the judge is asked ‘What is she being punished for?’ he must link his verdict with an event that can explain his decision. We cannot explain causes by their effects. I can explain the length of a shadow by the
length of the flagpole casting the shadow (plus the position of the sun). But I
cannot explain the length of the flagpole in terms of the length of the
shadow (even though I can predict the length of the flagpole from the length
of the shadow).

Perhaps the basis for the causal requirement in explanation is merely
pragmatic. When people ask why something happened, they want to know
the cause. They want to know the cause because that helps them manipulate
events. They want to manipulate events because that helps them achieve
their goals. Whatever the source, the causal requirement for explanation
suggests a constraint on the justification of verdicts. We cannot justify a
verdict by appealing to the crime it will cause.

4. The generality of the hunger for grounding

Colonel Henry Watterson was a newspaper editor during the late nineteenth
century era of American railroad travel. As a journalist, he and his reporters
were issued special railroad passes. These passes were non-transferable. This
restriction was widely violated. During one crackdown, a conductor con-
fronted a young man who presented the pass of a certain Mr. Smith who
was a correspondent for Watterson’s paper. The suspicious conductor told
the nervous man that Colonel Watterson happened to be on the same train.
He escorted “Mr. Smith” for authentication. The conductor asked,
“Mr. Watterson, is this man your employee?” To the young man’s amazement,
the answer was “Yes”. The conductor left. The relieved man began to
effusively express his gratitude. “Compose yourself, young man. I don’t
happen to be Colonel Watterson, but I am riding on his railroad pass.”

Narrowly construed, the practice of vouching for unknown parties only
transmits knowledge when the chain of vouches ends in a known voucher.
That’s why two men who are strangers to you cannot instill confidence by
vouching for each other.

However, we do get some independent information from the manner in
which the vouches are performed. Sometimes that is enough to make the
avowal rationally compelling. The ventriloquist Edgar Bergen was once
invited to the White House but was refused admittance because he had no
identification. Bergen searched all his pockets for suitable documents.
Finally, he opened the case containing his dummy, Charlie McCarthy, and
hoisted Charlie into view. Charlie McCarthy looked at Bergen, and then
looked at guards: “Yeah, fellows, he’s Edgar Bergen.” The ventriloquist was
allowed in on Charlie’s say-so. The moral is that some apparent counter-
examples to the grounding requirement conform to it when they are con-
strued to include collateral information.

Just as there are acceptable risks of uttering a falsehood, there are
acceptable risks of uttering ungrounded statements. Mr. Borrower can
safely contract to pay three percent more than the prime lending rate. For
the Federal Reserve does not specify banking rates parasitically as in ‘Three percent less than the mortgage rate paid by Mr. Borrower’. In contrast, the risk is too high for the practice of using sealed bids. Auctioneers require that bids be self-contained. One cannot bid ‘One dollar more than any of the other bids’. For if another person bids ‘One dollar more than any of the other bids’, the two bids would be ungrounded. Since the bidders are motivated to bid just slightly more than their competitors, these indirect bids would be popular and so threaten the decisiveness of the auction.

Despite what might be inferred from Christopher New’s kamikaze example, grounding is a significant concern with rewards. True, morality weights harms more heavily than benefits, so there is lower level of scrutiny for rewarding than punishing. Rewards are most apt to attract suspicion in competitive situations. Suppose that we predict that a driver will win a race because he has a much faster car. He has a much faster car because he could afford it with the help of the pre-reward money. The other drivers would object that pre-rewarding is viciously circular.

A driver can borrow from a banker on the strength of a prediction that the loan will cause him to win and thereby ensure repayment with interest. Here the belief in an upcoming reward causes the reward. However, the reward does not cause the victory. The reward is still an effect of the victory. We only object to the reward being a cause of the victory. I interpret this as a concern about ungroundedness.

Pre-forgiveness is also sensitive to the causal independence requirement. Suppose that the only consideration stopping my neighbor’s boy from smashing my Halloween pumpkin is fear that I will bear a grudge against him. Despite knowing this, I forgive him in advance for smashing the pumpkin. Such “forgiveness” seems tantamount to permission.

So why does it seem permissible to pre-reward kamikaze pilots for gallantry and to pre-forgive our executioners? Because we bend the rules. If I am guilt stricken after harming your baby, you may ease my self-recrimination by forgiving me on the baby’s behalf. But the only person who can forgive is the victim of the wrongdoing. One becomes a victim by experiencing the wrongdoing. Only then can you single out the wrongdoing you intend to forgive. The anticipation of a future pain can itself be painful and so one can forgive your torturer for your current dread. But forgiveness of the torture must await the torture itself.

When there is no opportunity to act after the fact, we turn a blind eye to infractions. Our impotence will also drive us to peculiar forms of post-punishment. After the Restoration, Oliver Cromwell was disinterred in 1661 from the tomb of kings in Westminster Abbey and hung from the gallows at Tyburn. This was presented as punishment of Cromwell for killing King Charles I. But I think it was too late to punish Cromwell. Similarly, I deny that Japanese commanders can reward their kamikaze pilots for future gallantry and I deny that Queen Mary forgave her
executioners for her upcoming decapitation. The futility of their gestures imbues their efforts with a poignancy that clouds the critical eye.

Complaints about grounding were first voiced by logicians hoping to diagnose the liar paradox. ‘This sentence is not true’ violates the requirement that the truth of a sentence be grounded within the real world. If sentence A attributes truth to sentence B, then the truth-value of A depends on the truth-value of B. The truth-value of sentence B may depend on a third sentence C. This is permitted as long as the chain ends in a finite number of steps and does not loop back on itself. Or so say the devotees of groundedness. Others complain that an unrestrained grounding requirement would deny a truth-value to logical laws such as ‘No proposition is both true and false’.

Various amendments to the grounding requirement have been proposed. These softenings carefully preserve condemnation of the truth-teller paradox ‘This sentence is true’. The amenders are also careful to condemn the looped truth-teller:

S. The next sentence is true.
L. The previous sentence is true.

My conjecture is that there are grounding requirements for decision-making. Perhaps these grounding requirements arise because decisions incorporate the concept of truth. When we decide, we resolve to bring about the truth of a certain description. Another possibility is that the grounding requirement enters by means of the belief component of decision-making. To believe a proposition is to believe it is true. We have trouble believing ‘This is true’ because we cannot get an independent grip on what is to be believed.

Hans Herzberger’s (1970, 150) used an infinite sequence to show that a chain can be ungrounded even without self-reference or referential loops.

1. The next sentence is true.
2. The next sentence is true.
3. The next sentence is true.
   .
   .
   .

This descending truth-teller serves as a model for demonstrating that a causal restriction is not sufficient to prevent all ungroundedness.

I am personally content with beginninglessness. I believe that the present could be the effect of an infinitely long causal chain and so believe that there could be an infinitely long chain of verdicts. Consider a prisoner who has an infinite past. Each year this recidivist commits the crime of trying to escape.
Each year a judge sentences him to another year in prison. Each verdict is caused by a crime. Yet there is no first crime in the sequence and hence no way to ground the whole chain of verdicts. Therefore, the causal requirement does not fully guarantee grounding.

Just as we want our uses of ‘true’ to eventually connect with facts (rather than yet other uses of ‘true’), we want our uses of ‘guilty’ to eventually connect with crimes (rather than yet other uses of ‘guilty’). Judges do not wish to reason in a concentrated circle. However, they are more tolerant of the dilute circles Nelson Goodman praised as “reflective equilibrium”.

A little ungroundedness may also be tolerated in an infinite scenario. To fully explain a daughter’s crime one may need to explain how her mother’s criminal convictions led to a bad upbringing for her daughter. Having introduced her mother into the story we may need to explain how her mother’s mother convictions led to a bad upbringing a generation earlier. One can imagine an infinite maternal lineage of convictions, each conviction being partly explained by an earlier conviction. Arguably, the verdict against the daughter would not be perfectly grounded. But this marginal ungroundedness is swamped by other elements of the story. There are infinitely many distinct agents receiving a variety of convictions from a variety of judges in diverse settings. Just as it is reasonable to accept a small risk of ungroundedness in making an assertion, it is acceptable for a judge to ignore ungroundedness when it plays an insignificant role in the explanation of the crime.

Significant injustice is generated in the case of the perpetual escapee because the scenario is homogenous: a single individual commits the same type of offense in the same setting. There can be significant injustice in a multiple agent scenario if the secondary nature of the offense is highlighted. Suppose Mr. One is convicted as an accessory after the fact for harboring the fugitive Mr. Two. Mr. Two was on the run because he was convicted as an accessory after the fact for harboring the fugitive Mr. Three. The chain of accessories goes back infinitely. Is the injustice merely due to remoteness? If so, then the injustice would not be decreased by eventually grounding the chain with a primary offender. But adding a base step to the recursion does help relieve some of the injustice. After all, if the authorities had been laboring under the false assumption that there was a primary offender (long, long ago), then the discovery that the chain stretches back infinitely would prompt the authorities to reconsider their continued prosecution of the accessories.

If a community became persuaded that there are infinite chains of punishment then they might develop safeguards against this source of ungroundedness. Since we believe that all these chains are actually finite, we have not erected any defenses.

Given that the causal requirement does not necessarily deliver groundedness, one wonders whether groundedness takes precedence over causation.
We get some illumination from R. M. Hare’s (1981) distinction between the intuitive and critical levels of moral thinking. At the intuitive level, we invoke rules, virtues, and motives. At the critical level, we go on to justify our promotion of these intuitive considerations. For instance, the utilitarian says we adopt rules that promote good consequences. Once the rules are adopted, we are generally content to condemn acts simply as infractions. But reformers can move to the critical level to change the rules.

Something similar can be said for causality and groundedness. I place the causal requirement for a just verdict at the intuitive level and the grounding justification at the critical level. A utilitarian might complain that groundedness is a rather opaque moral concern. Immanuel Kant might be less troubled because he often asserts a link between logic and morality (sometimes casting immorality as inconsistency). Hare is a utilitarian who is eager to co-opt Kant and his work provides some leads as to how a utilitarian might explain our concern with groundedness.

5. Time travel

Brian Weatherson has encouraged me to dramatize my position with a time travel scenario: After a law-abiding life, Sid suddenly turns vicious on his fortieth birthday. He commits a foul murder and flees back one year in his time machine. The police hop in their time machines and pursue Sid. When the police arrest him, Sid protests that his crime has yet to occur. The temporal ban on prepunishment supports Sid’s loophole. Sid is now free to commit another murder, and escape a further year into the past. Sid is on his way to becoming a serial killer!

The causal theory of verdicts nips this crime spree in the bud; a judge can convict Sid of his future crime since that crime is a (backward) cause of the verdict.

Sid has a more disturbing argument up his sleeve: Why not punish his thirty nine year old self as well? Since the forty year old caused the death and the forty year old is identical with the thirty nine year old, the thirty nine year old also caused the death. Yet the thirty nine year old is innocent. The problem is not that we would be doubling the amount of punishment; we would recoil from Sid’s suggestion that we flip a coin as to who gets punished or divide the punishment between the forty year old and the thirty nine year old.

This fanciful scenario suggests a metaphysical justification for prepunishment. If I deserve punishment at one stage of my life I also deserve it at any other stage of my life. After all, I am the same person through all my stages. Maybe this is how Christopher New reasons.

This argument does not use the causal theory of verdicts as a premise. It simply combines the principle that agents endure through time with temporal neutrality. Since one of my motivations for the causal theory is temporal
neutrality, I reject endurantism. I independently favor a four-dimensional view of time in which the agent perdures rather than endures. Sid is a space-time worm composed of stages that are not identical to each other. His forty year old stage is distinct from his thirty nine year old stage. One stage of Sid could deserve punishment without the other stage also deserving it. True, both stages would be causes of the death. We exonerate the thirty nine year old Sid in the same way as we exonerate Sid’s mother; both are causes but they are too remote to be appropriate causes.

6. Penal substitution

But why do we punish the forty one year old Sid for the crimes of the forty year old Sid? If a person stage can only be punished for the crimes it commits, then all punishment is too late. Guilty stages do not last long enough to be punished.

In extended sense, I espouse penal substitution (Lewis 1997). The familiar form of penal substitution involves the punishment of one person for the crime of another. In some societies, a son is punished for the father’s crime. We are tempted to dismiss this as absurd “punishment” of the innocent. However, Christians say Jesus atoned for the original sin of human beings through his crucifixion. And we all permit innocent third parties to pay heavy fines imposed on a convicted defendant. Plato and his wealthy friends wanted poor Socrates to save himself by offering to pay a large fine for his impiety. (In Athenian law of the era, the prosecution proposes one sentence while the convicted defendant proposes an alternative sentence. The jury chooses between the two sentences.)

Perdurantism suggests an extension of the concept of penal substitution. The idea is that we permit later person stages to be punished for the crime committed by an earlier stage. The criminal (as the sum of his person stages) is guilty by virtue of having a guilty part. But it does not follow that all of the ensuing person stages are guilty. True, they are all guilty in the top-down sense of being a person stage of a guilty person. But they are not thereby guilty in the bottom-up sense of being the stage that committed the crime. The objector who asks why we punish the post-crime Sid is thinking of guilt in this bottom-up way. From this unusual perspective, penal substitution is routine—even inescapable.

The causal and grounding requirements forbid punishment of pre-crime stages of Sid. Since these requirements are only necessary conditions for a just verdict, they only rule out punishments. To endorse punishments, we need sufficient conditions.

The classic rationales for punishment are poised to fill in the missing details. For instance, deterrence theorists can build on Derek Parfit’s (1984, 170–86) theme that people care much more about their future than their past. An amnesiac would prefer to learn that he just completed two hours of
judicial torture than that he is about to receive one hour of judicial torture. So when distributing the necessary evils, deterrence theorists will minimize evil by distributing the evil to future stages rather than past stages. Since we care more about our near future rather than our distant future, the deterrence theorist further recommends swift punishment. It is better to punish the forty one year old Sid than the sixty year old Sid.

Our future bias is neatly explained by evolutionary theory. We only have control over the future, so there will be natural selection in favor of agents who are preoccupied with what will happen.

7. Personal time v. external time

Does our concern with Sid’s stages show that the temporal asymmetry holds for “personal time”? Sid’s personal time is, roughly, the time as measured by his wristwatch (Lewis 1986, 69–70). It is a time-like relation that re-orders stages so that they fit together in the familiar way. In personal time, each stage of Sid is followed by an organically older stage (so his aging process would conform to physiology textbooks). In external time, some of Sid’s older stages (weirdly) precede his younger stages. Sid’s biography is apt to jump back and forth between a personal time line and an external time line. For instance, the biographer will write “Even now Sid may be stalking the streets of nineteenth century London.”

Since personal time is not time (Lewis 1986, 70), this is only a nominal incarnation of a temporal asymmetry. Furthermore, this re-sequencing is designed to preserve familiar causal patterns. So the promoter of personal time does not contest my thesis that causal prepunishment is necessarily unjust while temporal prepunishment is only contingently unjust.

He can agree that time, like money, cannot be intrinsically important. Generalizations that attribute agency to time must be exchanged for less abstract generalizations. ‘Time heals wounds’ cannot be accepted as a literal truth. Time is not efficacious. Old objects that seem marked by time are really marked by use or the weather or chemical processes. When a generalization about time looks appealing, we should look for an underlying generalization about causal processes.

8. Causal isolation

Brian Weatherson has raised a second issue about the time of killing. Suppose Mia has poisoned Woody’s cryogenic suspension. When Woody awakes from hyper-sleep in five hundred years, he will only last a week more. There is nothing to prevent this slow poisoning. Prosecutors accuse Mia of murder. Mia concedes the facts of the case but objects that Woody is not yet dead; she cannot be punished for a crime that has not yet taken place.
According to Donald Davidson (1969), killings occur as soon as the agent completes the appropriate bodily movement; we do not need to wait for the effects of that movement. If he is right, we do not prepunish when punishing Mia long before Woody dies. If you instead believe that the crime is only completed with Woody’s death, then you think part of the crime has transpired but not all of it. If the causal theory requires that the whole crime be a cause of the verdict then it forbids us from punishing Mia before Woody dies. If the causal theory only requires that part of the crime be a cause of the verdict, then it permits Mia to be punished.

There is much to be said for requiring only a part. Think of a crime that takes place over a wide space (or partly in the mind). Since our access is limited, our evidence will emanate from only some parts of the crime. We do not demand that our guilty verdict be caused by the whole crime. Of course, not any part will do. We want the “representative” parts employed in the causal theory of perception; I see the murderer if I see only his face but not if I only see his feet walking past my cellar window. The murderer’s act must be a representative part of his murder. So, if the punishment of Mia is prepunishment, then it is a species of prepunishment that the “representative part” variation of the causal theory does not forbid.

But what if Mia has an antidote and punishing her would stop her from saving Woody? Punishing Mia for murdering Woody would now be ungrounded: the punishment is causing its crime. This makes it tempting to wait for all the dust to settle before passing judgment.

There is no escaping frustration when it comes to punishment. Suppose Rocketman is about blast off on an exploration mission. Rocketman, Brian Weatherson explains, just wants to explore for his own sake and so does not have a radio. Rocketman blasts off and is never heard from again—just as expected. However, a security tape shows Ivy preparing to poison Rocketman’s air conditioning system. The tape runs out before showing her commencing the crime itself. Subsequent investigation produces overwhelming evidence that Ivy had the motive and ability to initiate the crime and complete it. No one reasonably doubts that she did it. But none of the evidence is an effect of any part of the crime. So the causal theory of verdicts clearly forbids us from convicting Ivy for the murder. Is the causal theory too fastidious?

No. Even when evidence from precursors gives us knowledge that the crime occurred, we continue to crave evidence that is an effect of the crime. If only we could get Ivy to confess or find a second security tape that shows some of the crime itself! It is such a shame that Rocketman carried no radio! If he did, then we could use the absence of a transmission as evidence that he was murdered. (Often when we appear to be relying solely on precursor evidence, we are relying on the causation of absences, for instance, the disappearance of the labor union leader Jimmy Hoffa, the lack of any fingerprints in the Hillside Strangler’s apartment, and, from the Sherlock
Holmes mystery “Silver Blaze”, the failure of the watch dog to bark when an expensive race horse was stolen from its stable.) The point of this pining for effect evidence is not that we wish to epistemically improve the belief that Ivy poisoned Rocketman. I have already stipulated that no one reasonably doubts that she did it. The point, I submit, is to get the right kind of pedigree for conviction. We put a premium on bodily remains and confessions because they are effects rather than mere precursors of the crime. The legal value of a broken clock at a crime scene is not exhausted by it indicating the time of the crime.

Antiquarians prefer to tell the time of day with antique clocks even though they are less reliable than modern clocks. Lawyers display a similar appetite for a causal connection (Thomson 1986). Statistical predictors are often recognized as being more reliable than effect evidence preferred by judges and juries. Some trade-off is rational. Is it just?

All of this lawyerly insistence on causal pedigrees competes with our other desire for a just outcome and good consequences. In cases of poetic justice and rough justice, we get a fitting result without the procedural burden. But there is no just punishment without a just process.

This is not Sunday school. I am not endorsing the proposition that we ought never to engage in unjust punishment. Organizers of the kangaroo trial of Nicolae and Elena Ceausescu during the 1989 Romanian Revolution explained that they needed a quick execution to undermine an imminent and catastrophic counter-revolution. Maybe they were right. Justice isn’t everything.

But justice is still something. Even hardheaded people undergo expense and inconvenience to secure it.

Acknowledgements

Earlier drafts were read at York University (UK), the University of Texas at Austin, the Syracuse Workshop in Metaphysics and Virginia Commonwealth University, and the University of Alabama. Especially helpful commentary was furnished by Brian Weatherson, Larry Crocker, and David Hershenov.

References


