

If legislators and judges all accepted the philosophical theory of determinism, what would be the effect on criminal sentencing?

Desert is the supposedly barbaric concept free will sceptics would have sentencing abandon, yet it is the only principle in criminal sentencing that calibrates coercive treatment to the moral gravity of the completed offence, while limiting the state's appeal to future protection. Determinism, on the incompatibilist reading this essay follows, would destroy it.¹

Desert performs two functions, the second dependent on the first: it provides the measure of punishment, the metric which tells a court what quantum fits the offence; and because it fixes that measure, it provides the limit, the principle that the state may not exceed what was deserved however useful a harsher sentence might be. The limit depends on the measure; destroy the measure, and the limit floats free.

In practice, a determinism-compatible system would lack the resources to distinguish between two defendants whose risk profiles are identical but whose crimes differ in moral gravity.

The four sentencing pillars (incapacitation, rehabilitation, deterrence, retribution) are entangled unevenly. Purely desert-based aggravating-factor enhancements collapse. Deterrence and incapacitation survive as causal mechanisms but lose their calibration and, with it, their ceiling: quantum scaled to offence severity, detention bounded by proportionality. Without desert anchoring those limits, the surviving pillars retain their purposes but not their boundaries.

Determinism alone does not collapse desert. Criminal law already operates on broadly compatibilist assumptions: voluntary action and mental capacity, not libertarian free will. The sentencing consequences follow only if legislators and judges accept the further step: determinism undermines the ultimate sourcehood basic desert (deserving suffering for its own sake, not for any further benefit)

requires.² A compatibilist court could preserve much of existing sentencing doctrine; the sharper question is what follows if legal actors accept determinism in the sourcehood-denying sense that places basic desert in doubt.

The strongest legal-compatibilist position is Morse's: criminal law requires only rational capacity and voluntary action, not ultimate sourcehood. Liability survives determinism on his terms: a defendant capable of practical reasoning remains criminally responsible. At sentencing, however, Morse defends retributive proportionality, endorsing Brink's 'analog responsibility' account. The capacity-vs-desert distinction the next section develops is where Morse's framework comes under pressure, and where both the metric and the limit it grounds would fail.³

If determinism is true and incompatibilism follows, no defendant is the originating cause of their crime. Frankfurt cases are widely taken to show alternative possibilities are neither necessary nor sufficient for moral responsibility. The difficulty is sourcehood: moral address survives determinism, practical accountability too.⁴ Basic-desert punishment does not. Without sourcehood, retributive justification collapses.

Pereboom's quarantine model is the alternative: forward-looking, rights-based, grounded in self-defence rather than moral blame.⁵ Incapacitation, rehabilitation, and deterrence survive as causal mechanisms. Retribution does not. But the quarantine analogy exposes the constraint problem: medical quarantine ends when the patient tests negative; criminogenic risk has no equivalent test, and the replacement inherits no natural limit on duration.

The distinction Morse blurs would prove fatal to his framework. Capacity is a threshold concept; desert is a calibration concept. Morse's rational-capacity criterion explains why the legal system can address a defendant. It does not explain why the sentence should be sized to what they deserve, or why the state must stop at

that size. If sourcehood fails, his proportionality criterion loses both its metric and its ceiling.

A court that accepts determinism can still decide who to sentence. It can no longer say by how much.

Neither capacity nor desert mechanically yields numerical quantum.⁶ Desert supplies the relevant dimension: the gravity of the completed wrong. Capacity can rank addressability; it cannot explain why months of imprisonment should track the seriousness of an offence.⁷

Limiting retributivism, the dominant model of US state guidelines⁸, positions desert as upper and lower limits, consequentialist purposes fine-tuning within. Without desert, these limits lose their anchor. Extending Bronsther's critique: the upper limit that protected defendants from disproportionate state coercion becomes an assertion without a principle.⁹ Frase acknowledges this: upper limits "reflect the widely shared sense that it is fundamentally unfair" to punish more than deserved.¹⁰ Remove desert, and that sense has nowhere to ground itself.

Desert anchors sentencing to the past act: the sentence responds to what was done, measured against its gravity. Without it, the question shifts from 'what did this person do?' to 'what is this person likely to do?', and the remaining justifications are all collective-interest. No remaining principle calibrates the individual defendant's sentence to what was done. Procedural rights constrain process; necessity and dignity cap coercion; but none ties quantum to the gravity of the completed offence.

Greene and Cohen (2004) argued neuroscience changes nothing about doctrine and everything about foundations.¹¹ The USSC's criminal-history axis confirms the entanglement: culpability, deterrence, and protection fused in a single metric.¹²

Mens rea, as legally relevant mental state, survives intact. What collapses is the inference from culpable mental state to deserved suffering.¹³

Pereboom's framework is defensible. Strawson's observation shows it is institutionally unstable when administered by human judges in the setting of criminal sentencing.

Pereboom's selective revision relies on the reflective distance possible in personal relationships (Nichols 2007); the sentencing courtroom is not such a context.

A judge ruling on a defendant before her cannot reliably sustain the reflective distance Pereboom's selective revision requires.¹⁴ The point is not that judges cannot maintain dispassionate composure; legal training is designed for that. The point is about outputs: reactive content shapes the discretionary range where guidelines stop and quantum is fixed.¹⁵

The reactive attitudes, resentment, indignation, and moral outrage, are constitutive of the participant stance, not rational positions one abandons. Strawson allows the objective attitude toward particular individuals: children, the mentally ill, the incapacitated. But universal determinism is not an exempting condition in his sense. Adopting the objective attitude toward everyone is what Strawson called practically inconceivable.¹⁶ I use Strawson's observation against his conclusion. The institutional persistence of reactive attitudes means human-administered sentencing cannot rest coercive outcomes on reasoning the state has rejected.¹⁷

A judge exercises not personal resentment but its vicarious analogue, what Strawson called "resentment on behalf of another, where one's own interest and dignity are not involved". He observed that suspending these vicarious attitudes universally is "perhaps even harder for us to envisage as a real possibility" than suspending the personal ones. A more careful objection: reactive attitudes ground informal censure, not proportionality principles for sentence quantum. That gap is

the problem. Judges import reactive content into a calibration task the content cannot perform.

Acceptance is an intellectual state; reactive attitudes operate independently. Legislators can design forward-looking schemes in the abstract but remain exposed through the electorate. Judges face the harder case: applying those schemes while looking a defendant in the eye.

The same reactive attitudes that contaminate forward-looking calibration also provide the system's last rough constraint: a judge who cannot suppress moral outrage at a brutal act will resist a disposition the risk assessment alone would permit. Irrationality is also a safeguard.

Once the system has rejected desert-based reasoning, tolerating variance correlated with rejected factors creates justificatory incoherence.¹⁸ The contamination has two components: reactive attitudes persist regardless of intellectual conviction, and without desert, no replacement preserves its dual function. Capacity survives as a threshold, but the mechanism that simultaneously fixed quantum and capped the state's authority has no substitute.¹⁹

Consider the USSC guidelines grid. The criminal-history axis currently serves culpability, deterrence, and protection simultaneously. Remove desert, and the culpability component loses its justification, but the grid cannot simply delete it because quantum calculations depend on it. Sentencing commissions would need to rebuild calibration on purely forward-looking grounds, but no forward-looking principle tells them whether a Category III offender should receive 18 months or 36. The grid becomes a structure without a metric.

State v. Loomis (2016) shows the gradient and the resistance at once.²⁰ The Wisconsin Supreme Court approved limited COMPAS (a risk-assessment tool) use under a five-part written advisement, forbidding it as the determinative factor. The court preserved human discretion, where individual circumstances still register but

rejected desert intuitions return unchecked. Resistance is evidence not that the pressure is absent but of what the system is not yet willing to lose.

Risk instruments trained on historical data inherit retributive biases, but the biases are at least inspectable in ways human discretion is not. Chouldechova and Kleinberg have shown the deeper problem: no prediction system can satisfy competing fairness metrics across different base rates. Algorithms do not solve calibration. They relocate it, and in doing so, remove the human reluctance that informally constrained it.

Caruso's quarantine model incapacitates the dangerous on public-health grounds. The model supplies a non-desert justification for confinement, and Pereboom's Principle of Least Infringement gives it a ceiling, no more restriction than necessary. But a ceiling is a constraint, not a metric. Quarantine tells us why restriction is permitted and when it goes too far. It does not tell us how much is enough. The limit survives; the measure does not.²¹ Caruso exposes the bind.

Legislators would face their own version of the pressure. Statutes would be rewritten toward public protection, rehabilitation, and risk management. Mandatory minimums and desert-based aggravating factors would become harder to justify. Sentencing commissions would gain importance. But legislators would also discover the constraint gap: statutes can mandate reasons and regulate risk instruments, but cannot replace the principle that told judges when enough was enough.

Desert-based aggravating factors persist unreplaced. Mens rea hierarchies can invert: where impulsivity predicts recidivism better than premeditation, reckless actors score as more dangerous than purposeful ones. Proportionality limits become unstable: everyone feels they should exist, but no one can say where they sit. Deterrence, freed from its hidden retributive limiter, licenses harsh sentences scaled to expected deterrent yield.²² The history confirms the prediction:

indeterminate sentencing regimes grounded in rehabilitative ideals did not deliver the leniency that rationale promised.²³

Consider two defendants convicted of serious assault. Same risk profiles, negligible. Both have completed structured rehabilitation. The only remaining difference is the moral gravity of the original act: A's assault was premeditated and brutal, B's impulsive and provoked. With risk and rehabilitation equalised, forward-looking calibration produces identical dispositions. Desert does not. If the system still treats them differently, it can only be on desert grounds the system has rejected. If it does not, sentencing has become morally alien.²⁴ That alienness is not a counterargument. It is the predicted incoherence at the level of intuition.²⁵

The system would not choose cleanly. It would default to human judges continuing to sentence while knowing the justificatory basis has been voided, retaining desert-shaped reasoning while officially denying desert. The effect is the gap itself: Smilansky's tragic illusionism (maintaining free-will belief despite knowing it false) is this prediction stated honestly.

A system built from scratch on forward-looking grounds might never have to confront the gap. Retributively saturated common-law systems do.

Three pressures, none resolvable. Keep human judges and accept contamination: the worst calibrators are also the last protectors, reactive attitudes distort quantum but resist dispositions permitted by pure risk logic.²⁶ Adopt algorithmic determination and accept depersonalisation: cleaner calibration, but no human reluctance standing between the defendant and the state's demand for security. Abandon determinism-compatible justice and accept incoherence: the state coerces defendants on grounds it has acknowledged as empty. The system would inhabit all three, defaulting to the first because it is least dangerous to the individual defendant, the second creeping at the margins, the third describing what the first is. Incoherence is not a failure mode. It is the steady state.

The system would discover the concept it had discarded as metaphysical confusion was also its most effective constraint, and the mechanism that held the individual defendant's liberty against the collective demand for protection.

Determinism would not abolish sentencing. It would deprive sentencing of the concept that tells punishment how much is enough. The institutions would continue, the judges would continue, the sentences would continue; the philosophy that once told them how to stop would not.

Endnotes

1. The framing of desert as the only principle that calibrates coercive treatment to the moral gravity of the completed offence has its lineage in Hart's insistence that the distribution of punishment must respect individual fairness regardless of the general justifying aim (1968), and in Locke's political philosophy of constraining sovereign power. Other principles, dignity, parsimony, necessity, can cap coercion; the specific claim that only desert ties quantum to offence gravity is the argument advanced here.
2. Locke's locked-room thought experiment (Essay II.xxi.10) anticipates Frankfurt cases by dissociating voluntariness from the ability to do otherwise: a man who remains in a locked room because he wishes to converse with his friend acts voluntarily even though he could not have left. The voluntary/free distinction maps onto modern criminal law's actus reus threshold. Fischer and Ravizza's reasons-responsiveness account, which preserves desert without ultimate sourcehood, descends from Locke's analysis of the power to suspend desire at II.xxi.47-48; Yaffe, by contrast, reads II.xxi.10 as consistent with incompatibilism. This essay acknowledges the compatibilist lineage but follows the prompt's distinctive force.
3. A 333-year continuity runs from Locke II.xxi.14, where the question whether the will is free is "altogether improper", to Morse's 2023 claim that free will "in the strong, metaphysical sense is not a criterion for any criminal-law doctrine". Both deny that criminal law requires libertarian free will. This essay accepts the continuity at liability and pivots at sentencing: even on the strongest legal-compatibilist reading, retributive proportionality imports a desert calibration, and the constraint derived from it, that the position cannot ground without smuggling back the metaphysics it claims to reject.
4. Duff (2001) develops the leading communicative theory of punishment, on which hard treatment is justified as a mode of moral dialogue: the state addresses the offender as a responsible agent, communicates censure for the wrong, and invites repentance and reconciliation. The theory's status under determinism turns on what grounds the censure. If censure tracks backward-looking moral judgment about what the offender deserves to hear, it imports the desert basis the previous section showed determinism removes, and collapses with retribution. It also inherits the constraint problem: without desert fixing how much censure the wrong warrants, communicative intervention has no principled endpoint. If censure is reconstructed as forward-looking, aimed at moral formation, reintegration, or the expressive functions Feinberg (1965) identified, it survives the philosophical collapse but inherits the Strawsonian problem developed later: the communicative posture requires treating the defendant as a participant in moral dialogue, which is the participant stance the essay argues human judges cannot coherently abandon. Either way, communicative theory does not offer a third path between the collapse and the human hurdle. It faces both in sequence.
5. Pereboom's framework imposes three explicit constraints: less restrictive measures must be tried first, the state has an obligation to rehabilitate, and no greater harm may be inflicted than is necessary for protection. The right-to-harm grounding rules out using individuals merely as means in ways pure consequentialism would permit. The framework also operates post-conviction. Actus reus remains the threshold for state intervention; the probabilistic element enters only at calibration of response. Innocent people cannot be incapacitated at all because they pose no threat. This is why the model is properly described as forward-looking and rights-based rather than consequentialist: the right-to-harm constraint forecloses the welfare trade-offs a purely consequentialist justification would permit. These rights-based constraints are the non-desert ceiling the body identifies as surviving the loss of desert; what they cannot supply is the metric beneath it. Morse (2023) is the direct interlocutor. He calls the public-health-quarantine model "a dystopian regime that neither Brink nor I could endure", aligning with Brink and Sehon in rejecting hard incompatibilism.
6. The claim is not that desert yields a precise numerical formula. It is that desert provides an offence-centred dimension along which legislation, precedent, and convention fix numerical quantum. Without that dimension, the system must anchor quantum to risk or necessity, which have no proportionality link to the gravity of the past act.
7. The US Supreme Court has spent two decades calibrating sentence severity by capacity: *Atkins v Virginia* (2002), *Roper v Simmons* (2005), *Graham v Florida* (2010), *Miller v Alabama* (2012). The cases produce categorical exclusions, not continuous metrics. They tell us no juvenile gets the death penalty; they do not tell us a 19-year-old of diminished capacity gets fifteen years rather than twenty. Threshold, not calibration.
8. This essay focuses on common-law systems, using U.S. sentencing structures as its primary illustration. The argument generalises to any retributively saturated system, but the institutional machinery cited is predominantly American. The Frase 2004 "consensus model" formulation builds on Norval Morris (1974).
9. The structural argument has its lineage in Hart's mixed theory (1968), Norval Morris's (1974) "limiting retributivism", and Frase's (2004) "consensus model" describing the dominant pattern in US state guidelines. The position taken here radicalises the observation: retribution is structurally parasitic on the surviving pillars rather than a discrete fourth purpose operating alongside them. The further claim, that the constraint depends on the calibration and that losing the metric necessarily destabilises the ceiling, extends beyond existing literature: Morris identified desert's limiting function and Hart separated justification from distribution, but neither argued the dependency. Frase himself concedes the point in foundational terms, acknowledging that upper limits "reflect the widely shared sense that it is fundamentally unfair and an abuse of governmental power to punish an offender more severely than he deserves". Once desert is rejected as a justificatory ground, the sentiment Frase invokes loses its anchor. Bronsther (2021) presses the same structural point, arguing that limiting retributivism's upper limits cannot do justificatory work without basic desert. Tadros (2011) defends a duty view on which wrongdoers acquire enforceable duties to bear punishment, grounding proportionality without basic desert. The duty view still requires a metric for how much duty the wrongdoer owes; that metric is doing desert's work under a different name. The position is contested (Ferzan, Bergelson, Lippert-Rasmussen).
10. The US Sentencing Commission's hybrid grid illustrates the entanglement at the doctrinal level. The criminal-history axis serves culpability and recidivism together by the Commission's own framing, while the offence-level axis dominates the final quantum.

Mandatory minimums function as desert-mimicking calibration even where the stated rationale is deterrent (Tonry; von Hirsch; Frase). *R v Smith* [1987] 1 SCR 1045 makes the point in litigated form. The Canadian Supreme Court struck down the seven-year mandatory minimum for importing any quantity of narcotics as grossly disproportionate to its stated deterrent purpose. The retributive-structure reading is the argument advanced here rather than the Court's: whatever rationale policymakers offered, the statute's structure, with quantum fixed by offence category rather than individual risk, operated as desert-mimicking calibration in Frase's sense.

11. Greene and Cohen's philosophical analysis, the distinction between doctrine and its justificatory foundations, remains compelling even though their broader neuroscience prediction has not materialised. Weisberg et al. (2008) on the "seductive allure" of neuroscientific explanation supports the partial diagnosis, though the larger forecast has not borne out; Morse (2023) calls neuroscience-scepticism "the current craze". Strawson's phenomenological argument converges with Greene and Cohen's cognitive-systems account. Folk psychology, the gateway to moral evaluation, treats agents as uncaused causers, while folk physics treats mechanisms causally. The participant stance is structurally entrenched, not merely habitual.

12. Robinson (2013) argues tracking community desert judgments sustains compliance, supplying a consequentialist justification for retributive practice. Three responses bear on the structural argument here. First, community desert judgments are where racial and class biases concentrate, and minority perspectives are systematically underrepresented in the surveys these judgments rest on (Braman, Kahan & Hoffman 2010; Slobogin 2013; Ristroph). Enshrining these intuitions as policy consecrates discriminatory outcomes. Second, even granting Robinson's legitimacy claim, what it justifies is retributive communication, not retributive calibration; community intuitions were themselves shaped by the retributive system, so the appeal is circular. Third, Slobogin's empirical work shows community agreement is much weaker on appropriate punishments than on ordinal ranking of crimes, undermining the calibration use to which Robinson puts the data.

13. *Mens rea* operates throughout criminal law as a graded hierarchy of culpable mental states (purpose, knowledge, recklessness, negligence) running across thousands of distinct offences. That hierarchy survives determinism: a defendant who acts with purpose relates to the prohibited result differently from one who acts negligently, and the legal system can register that difference without committing to ultimate sourcehood. What collapses is the further inference Morse defends, that gradations in mental state should calibrate gradations in deserved suffering. The hierarchy can sort defendants by capacity. It cannot, on its own, fix the quantum of punishment owed. Morse's trajectory across two decades runs in the opposite direction. His earlier work defended generic partial-responsibility doctrine (2003); his 2023 review essay endorses Brink's trivalent and tetravalent culpability assessments, calibrating sentence to ever-finer gradations of analog responsibility. The institutional response to determinism described here would require movement the other way: away from finer retributive calibration, toward forward-looking criteria the surviving pillars can actually ground. A note on insanity. Expanding the insanity defence addresses liability rather than sentencing calibration; expanding sentencing-stage exemptions runs into the calibration problem identified here. Neither rescues retributive proportionality from the loss of sourcehood.

14. Pereboom argues the reactive attitudes can be selectively revised, with moral concern, sorrow, and resolve substituting for resentment and indignation without loss of what genuinely matters in personal relationships. The narrower claim of this essay is that institutional sentencing blocks even this selective revision, because the courtroom is not a context that permits the wide-profile reflective standing-back Pereboom's argument relies on (Nichols 2007). The substitutes function in personal relationships partly because the participants can step out of the immediate exchange to reflect on their own attitudes; a sentencing judge, ruling in real time on a defendant before her, cannot.

15. Lacey and Pickard (2013) argue the courtroom can be reformed for detached accountability without affective blame. Problem-solving courts, drug courts, mental-health courts, juvenile courts, demonstrate the possibility in non-adversarial, consent-based, low-stakes contexts. The argument here is narrower: coercive adversarial sentencing of serious crimes, where the bind operates.

16. Moore is the paradigmatic desert-based retributivist; he grounds desert in objective wrongness rather than alternative possibilities, which sidesteps Frankfurt cases but not the sourcehood objection developed above. Smilansky's illusionism (2000) defends maintaining belief in libertarian free will despite knowing the belief is illusory, on the ground that civilisation rests on it. He calls the defence tragic, the illusion "repugnant and demeaning" but unavoidable. His 2017 critique of Pereboom argues that hard incompatibilist alternatives like quarantine collapse into "funishment": detention either too comfortable to deter, or harsh enough to reintroduce the desert basis it claims to abandon. The body returns to this below. Emilsson (2026) re-read Strawson's inescapability claim as metaphysical rather than psychological. The bind works on either reading. If psychological, the empirical record on judicial discretion confirms it; if constitutive, the bind sharpens, because the problem is not remediable even in principle. Watson (1987, 2014) distinguishes attributability from accountability and argues compassionate accountability is possible. Exceptional individuals demonstrate this; the institutional question is whether thousands of judges across daily dockets can sustain it. The argument here predicts they cannot. Vargas (2013) defends moderate revisionism as a third path. Its target is responsibility writ large, not coercive sentencing specifically.

17. The persistence of reactive attitudes is well-established (Watson 1987, Pereboom 1995, Wallace 1994). The argument here is the institutional consequence framed as a bind, not the persistence claim itself.

18. Judicial debiasing research indicates that judges harbour implicit biases that can influence judgment but can be compensated for given sufficient motivation (Rachlinski and Johnson 2009); training reduces but does not eliminate extra-legal factors in sentencing. Once the system has accepted desert-based reasoning is invalid, knowingly allowing rejected reasoning to shape outcomes is tolerating injustice. Parole boards illustrate the hurdle in operation: nominally risk-based decisions are shaped by subjective remorse assessments that boards cannot reliably interpret (Young and Chimowitz 2022). A further worry runs the other way: determinism may harden sentencing rather than soften it, replacing "you deserve punishment" with "you are permanently dangerous". The most historically plausible institutional response, reclassifying free will as a legal fiction in the manner of corporate personhood, fails because the fiction would underwrite coercive state power on a basis the system has rejected. Strawson himself observed that

"thoroughgoing objectivity of attitude excludes essential elements in the concepts of moral condemnation and moral responsibility". The legal fiction asks the system to maintain those concepts in form while excluding their substance, which is not pragmatic fiction but knowing injustice. Legitimacy-based compliance (Tyler 1990) presupposes a public that believes sentences track desert; a system openly coercing on grounds it concedes are empty forfeits that legitimacy. Dignity-based limits can survive independently: *Vinter v United Kingdom* (2013) held whole-life terms without review to violate Article 3, illustrating a ceiling grounded in dignity rather than desert.

19. McGeer's "Civilizing Blame" (2013) argues that reactive attitudes can be scaffolded toward constructive ends through institutional design. In non-coercive settings such as restorative conferences and therapeutic communities, this may succeed; McGeer herself invokes restorative justice and concedes that current blame practices produce discriminatory outcomes. The scaffolding view suggests a fourth path beyond the trilemma developed below: institutional reform toward dialogical, restorative practice rather than retributive punishment or algorithmic determination. Restorative justice operates well for low-stakes offences with willing participants. It does not scale to coercive sentencing of serious crimes against unwilling defendants. Within that context, McGeer's view collapses into the first horn of the trilemma: reactive attitudes are retained because they cannot be eliminated, and their retention generates the legitimacy problem this section describes. The restorative alternative is a genuine fourth path for non-coercive contexts; the argument here brackets it as outside the scope of coercive sentencing.

20. *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), cert. denied 137 S. Ct. 2290 (2017). The Wisconsin Supreme Court's five-part written advisement on COMPAS use require sentencing courts to be informed that: the proprietary algorithm prevents disclosure of how factors are weighted; scores are based on group rather than individual data; the tool was developed for non-sentencing populations; studies have raised questions about racial disparity; and the instrument must be regularly monitored for accuracy. The court forbade COMPAS as the determinative factor in sentencing. *Loomis* is not a precedent for algorithmic sentencing but a guardrail against it; the algorithmic endpoint described here requires moving past the limit even sympathetic courts have placed on algorithmic sentencing. Mayson (2019) develops the "bias in, bias out" critique, distinguishing between predictive accuracy and normative permissibility of input variables. Some predictively useful variables are normatively impermissible regardless of their accuracy. Three-strikes regimes anticipate the multiplier effect at issue. Isaac Taylor (2023) argues from condemnation that meaningful public control requires sentencing decisions to remain under the control of an agent with standing to act on behalf of the wider political community; the algorithmic endpoint either violates this constraint or has to be paired with human ratification that reintroduces the contamination identified above. Danaher (2022) argues societies delegate tragic moral choices to autonomous systems to absorb the psychological burden. Algorithmic sentencing fits the pattern: the limit case of forward-looking calibration also functions as the institutional escape valve from the choice the system can no longer ground.

21. Desert-based sentencing satisfies Fuller's formal conditions of legality, prospectivity, generality, congruence between declared rule and official action, in ways that risk-based disposition does not. "Proportionate to the act" is a rule applicable in advance; "reasonable given the risk" is a standard applied after the fact. The dual loss is also a shift in the kind of legal constraint the system operates under.

22. Wootton (1963) followed determinism to its administrative conclusion: courts 'free to deal with every lawbreaker in whatever way best calculated to discourage future lawbreaking.' No proportionality constraint, no backward-looking anchor, no upper bound. The programme this essay predicts is Wootton's, with the dual-loss diagnosis this essay supplies. General-prevention theory establishes deterrence as a freestanding sentencing rationale (Andenaes 1952, 1966); nothing in determinism removes its predictive logic, only the desert ceiling that formerly bounded it.

23. The history of indeterminate sentencing in the United States is documented in Tonry, *Sentencing Matters* (1996). Indeterminate, treatment-oriented regimes did not deliver the leniency their rehabilitative rationale promised. The Sentencing Reform Act of 1984, which introduced federal sentencing guidelines, was motivated in significant part by the recognition that indeterminate sentencing had failed to produce either consistency or leniency (Stith and Cabranes 1998). The constraint-loss reading is the argument advanced here: without desert as ceiling, rehabilitative detention expanded to fill the space the ceiling had bounded.

24. A consequentialist system would address victims' legitimate needs for acknowledgment and public recognition of wrong through non-retributive mechanisms: restorative conferences, public findings, structured reparations. The framing of "closure" is contested in trauma literature and is not what the institutional response would be aiming at. In the assault scenario above, both victims retain an interest in public recognition of the wrong done. What determinism removes is the inference from that recognition to continued hard treatment of an offender no longer dangerous. The interest in acknowledgment survives the loss of the desert metric; the interest in proportional suffering does not.

25. A determinist who accepts the morally-alien result can maintain a principled forward-looking ceiling through Pereboom's Principle of Least Infringement: no more restriction than necessary for protection. That ceiling is genuine and individual-regarding. What it cannot do is perform the offence-centred calibration desert performed. It tells the state when to stop coercing; it does not tell the state why this offender's sentence should differ from another's whose risk profile is identical but whose crime was graver. The ceiling survives; the dimension of judgment beneath it does not.

26. Reactive attitudes function in three registers across the argument. Earlier they are contamination resisting elimination. In the same passage they are also the system's last rough constraint: the outrage that distorts calibration also resists dispositions pure risk logic would permit. In the trilemma below they are constitutive of treating defendants as persons rather than risk profiles. The bind is that all three are the same phenomenon.

Bibliography

- Andenaes, Johannes. "General Prevention—Illusion or Reality?" *Journal of Criminal Law, Criminology, and Police Science* 43, no. 2 (1952): 176–198.
- Andenaes, Johannes. "The General Preventive Effects of Punishment." *University of Pennsylvania Law Review* 114, no. 7 (1966): 949–983.
- Atkins v. Virginia*, 536 U.S. 304 (2002).
- Bergelson, Vera. "The Duty to Protect the Victim – Or the Duty to Suffer Punishment?" *Law and Philosophy* 32, no. 2–3 (2013): 199–215.
- Braman, Donald, Dan M. Kahan, David A. Hoffman, Paul H. Robinson, Owen D. Jones, Robert Kurzban, and John Darley. "Some Realism about Punishment Naturalism (with Responses)." *University of Chicago Law Review* 77 (2010): 1531–1660.
- Brink, David O. *Fair Opportunity and Responsibility*. Oxford: Oxford University Press, 2021.
- Bronsther, Jacob. "The Limits of Retributivism." *New Criminal Law Review* 24, no. 3 (2021): 301–325.
- Caruso, Gregg D. *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*. Cambridge: Cambridge University Press, 2021.
- Caruso, Gregg D. "Skepticism About Moral Responsibility." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Stanford University, Spring 2021 ed. <https://plato.stanford.edu/archives/spr2021/entries/skepticism-moral-responsibility/>.
- Chouldechova, Alexandra. "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments." *Big Data* 5, no. 2 (2017): 153–163.
- Danaher, John. "Tragic Choices and the Virtue of Techno-Responsibility Gaps." *Philosophy & Technology* 35, no. 2 (2022): article 26.
- Duff, R. A. *Punishment, Communication, and Community*. New York: Oxford University Press, 2001.
- Emilsson, Anton. "Strawson's Inescapability Claim: Metaphysical, not Psychological." *Journal of the American Philosophical Association* 12, no. 1 (2026): 110–128. Published online October 8, 2025. doi:10.1017/apa.2025.10015.
- Feinberg, Joel. "The Expressive Function of Punishment." *The Monist* 49, no. 3 (1965): 397–423.
- Ferzan, Kimberly Kessler. "Rethinking The Ends of Harm." *Law and Philosophy* 32, no. 2–3 (2013): 177–198.
- Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 1998.
- Frankfurt, Harry G. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66, no. 23 (1969): 829–839.
- Frase, Richard S. "Limiting Retributivism." In *The Future of Imprisonment*, edited by Michael Tonry, 83–119. New York: Oxford University Press, 2004.
- Frase, Richard S. "Punishment Purposes." *Stanford Law Review* 58, no. 1 (2005): 67–83.
- Fuller, Lon L. *The Morality of Law*. Rev. ed. New Haven: Yale University Press, 1969.
- Graham v. Florida*, 560 U.S. 48 (2010).
- Greene, Joshua, and Jonathan Cohen. "For the Law, Neuroscience Changes Nothing and Everything." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359, no. 1451 (2004): 1775–1785.
- Hart, H. L. A. *Punishment and Responsibility: Essays in the Philosophy of Law*. Oxford: Clarendon Press, 1968.
- Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan. "Inherent Trade-Offs in the Fair Determination of Risk Scores." In *Proceedings of the 8th Conference on Innovations in Theoretical Computer Science (ITCS 2017)*, edited by Christos H. Papadimitriou, vol. 67, 43:1–43:23. Dagstuhl: Leibniz International Proceedings in Informatics, 2017.
- Lacey, Nicola, and Hanna Pickard. "From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm." *Oxford Journal of Legal Studies* 33, no. 1 (2013): 1–29.
- Lippert-Rasmussen, Kasper. "'To Serve and Protect': The Ends of Harm by Victor Tadros." *Criminal Law and Philosophy* 9, no. 1 (2015): 49–71.
- Locke, John. *An Essay Concerning Human Understanding*. Edited by Peter H. Nidditch. Oxford: Clarendon Press, 1975. First published 1690.
- Loomis v. Wisconsin*, 137 S. Ct. 2290 (2017) (cert. denied).
- Mayson, Sandra G. "Bias In, Bias Out." *Yale Law Journal* 128, no. 8 (2019): 2218–2300.
- McGeer, Victoria. "Civilizing Blame." In *Blame: Its Nature and Norms*, edited by D. Justin Coates and Neal A. Tognazzini, 162–188. New York: Oxford University Press, 2013.
- McGeer, Victoria. "P. F. Strawson's Consequentialism." In *Oxford Studies in Agency and Responsibility*, Volume 2: 'Freedom and Resentment' at 50, edited by David Shoemaker and Neal Tognazzini, 64–92. Oxford: Oxford University Press, 2014.
- Miller v. Alabama*, 567 U.S. 460 (2012).
- Moore, Michael S. *Placing Blame: A Theory of the Criminal Law*. Oxford: Clarendon Press, 1997.

- Morris, Norval. *The Future of Imprisonment*. Chicago: University of Chicago Press, 1974.
- Morse, Stephen J. "Criminal Responsibility Reconsidered." *Criminal Law and Philosophy*, published online August 30, 2023. Print version: 19, no. 2 (2025): 315–329.
- Morse, Stephen J. "Diminished Rationality, Diminished Responsibility." *Ohio State Journal of Criminal Law* 1, no. 1 (2003): 289–308.
- Nichols, Shaun. "After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes." *Philosophical Perspectives* 21, no. 1 (2007): 405–428.
- Pereboom, Derk. "Determinism al Dente." *Noûs* 29, no. 1 (1995): 21–45.
- Pereboom, Derk. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press, 2014.
- Pereboom, Derk. *Living Without Free Will*. Cambridge: Cambridge University Press, 2001.
- R v Smith*, [1987] 1 S.C.R. 1045 (Supreme Court of Canada).
- Rachlinski, Jeffrey J., and Sheri Lynn Johnson. "Does Unconscious Racial Bias Affect Trial Judges?" *Notre Dame Law Review* 84, no. 3 (2009): 1195–1246.
- Ristroph, Alice. "Proportionality as a Principle of Limited Government." *Duke Law Journal* 55, no. 2 (2005): 263–331.
- Robinson, Paul H. *Intuitions of Justice and the Utility of Desert*. New York: Oxford University Press, 2013.
- Robinson, Paul H., and John M. Darley. "Intuitions of Justice: Implications for Criminal Law and Justice Policy." *Southern California Law Review* 81, no. 1 (2007): 1–67.
- Roper v. Simmons*, 543 U.S. 551 (2005).
- Sehon, Scott. *Free Will and Action Explanation: A Non-Causal, Compatibilist Account*. Oxford: Oxford University Press, 2016.
- Slobogin, Christopher, and Lauren Brinkley-Rubinstein. "Putting Desert in Its Place." *Stanford Law Review* 65, no. 1 (2013): 77–135.
- Smilansky, Saul. *Free Will and Illusion*. Oxford: Clarendon Press, 2000.
- Smilansky, Saul. "Pereboom on Punishment: Funishment, Innocence, Motivation, and Other Difficulties." *Criminal Law and Philosophy* 11, no. 3 (2017): 591–603.
- State v. Loomis*, 881 N.W.2d 749 (Wis. 2016).
- Stith, Kate, and Jose A. Cabranes. *Fear of Judging: Sentencing Guidelines in the Federal Courts*. Chicago: University of Chicago Press, 1998.
- Strawson, P. F. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1962): 1–25. Reprinted in *Free Will*, edited by Gary Watson, 2nd ed., 72–93. Oxford: Oxford University Press, 2003.
- Tadros, Victor. *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press, 2011.
- Taylor, Isaac. "Justice by Algorithm: The Limits of AI in Criminal Sentencing." *Criminal Justice Ethics* 42, no. 3 (2023): 193–213.
- Tonry, Michael. *Sentencing Matters*. New York: Oxford University Press, 1996.
- Tyler, Tom R. *Why People Obey the Law*. New Haven: Yale University Press, 1990.
- United States Sentencing Commission. *Demographic Differences in Federal Sentencing*. Washington, DC: U.S. Sentencing Commission, 2023.
- United States Sentencing Commission. *Measuring Recidivism: The Criminal History Computation of the Federal Sentencing Guidelines*. Washington, DC: U.S. Sentencing Commission, 2004.
- Vargas, Manuel. *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press, 2013.
- Vinter and Others v. United Kingdom*, App. Nos. 66069/09, 130/10, and 3896/10 (Eur. Ct. H.R., July 9, 2013).
- Von Hirsch, Andrew. *Censure and Sanctions*. Oxford: Clarendon Press, 1993.
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.
- Watson, Gary. "Peter Strawson on Responsibility and Sociality." In *Oxford Studies in Agency and Responsibility*, Volume 2: 'Freedom and Resentment' at 50, edited by David Shoemaker and Neal Tognazzini, 15–32. Oxford: Oxford University Press, 2014.
- Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by Ferdinand Schoeman, 256–286. Cambridge: Cambridge University Press, 1987.
- Weisberg, Deena Skolnick, Frank C. Keil, Joshua Goodstein, Elizabeth Rawson, and Jeremy R. Gray. "The Seductive Allure of Neuroscience Explanations." *Journal of Cognitive Neuroscience* 20, no. 3 (2008): 470–477.
- Wootton, Barbara. *Crime and the Criminal Law: Reflections of a Magistrate and Social Scientist*. London: Stevens, 1963.
- Yaffe, Gideon. *Liberty Worth the Name: Locke on Free Agency*. Princeton: Princeton University Press, 2000.
- Young, Kathryn M., and Hannah Chimowitz. "How Parole Boards Judge Remorse: Relational Legal Consciousness and the Reproduction of Carceral Logic." *Law & Society Review* 56, no. 2 (2022): 237–260. DOI: 10.1111/lasr.12601.