

Retributivist Arguments against Presuming Innocence

Answering to Duff

Alwin A. van Dijk

1 Introduction

I begin with arguing for an unusually broad PoI: anything worth discussing counts as a PoI interference. The aim of this PoI is to structure the debate about the PoI around pro- and anti-PoI values. I use Duff's retributivist punishment theory to generate such values.

Subsequently, I analyse three procedures that interfere with the PoI. Retributivists, like Duff, tend to castigate such procedures based on their (supposed) consequentialist rationale. I argue that they might be provided with a retributivist rationale as well.

2 Plea for a Broad PoI

Determining the scope of a PoI is a daunting task. The PoI of the European Convention on Human Rights (ECHR) appears to be a rather straightforward command when read in isolation: everyone charged with a criminal offence shall be presumed innocent until proved guilty according to law. But a naïve observer – such as I – is in for a huge surprise when he explores the comprehensive and complicated case law. Many instances in which innocence is clearly not presumed are nonetheless judged not to violate the PoI. Pre-trial detention, presumptions of guilt and inferring guilt from silence may all be compatible with this PoI!

The term *violation* is used cautiously in the ECHR's context, because there are grave legal consequences attached to it. Thus, it must only be used if something is wrong altogether. If pre-trial detention is deemed legitimate to secure interests outside the PoI, this should not be classified as a PoI violation. This dichotomic framework – to violate or not to violate – has led to a confusing academic debate.¹ A legal phenomenon is either labelled as consistent or as inconsistent with one and the same PoI. It is often unclear whether this is due to a different interpretation of the PoI or a different weighing of other values.

Asserting that something is a PoI violation is not a neutral act, and might have the same rhetorical dynamics as the *argumentum ad Hitlerum*. The association

1 See Loncke Stevens, 'The Meaning of the Presumption of Innocence for Pre-trial Detention. An Empirical Approach,' this issue, s. 2 and 4.

Alwin A. van Dijk

with something odious, like a PoI violation, puts the opponent in a difficult position. He now has to argue that something in which innocence is clearly not presumed is nonetheless compatible with the presumption of innocence. In my view, the discussion may benefit from using a different terminology. If all themes under discussion are classified according to what they really are – interferences with a PoI of some sort – the playing field will become more even. This idea builds on Duff's fruitful suggestion that we need not restrict ourselves to analysing *the* PoI. But where Duff leaves room for many PoI in different contexts, my PoI remains constant in all contexts.

I propose the following negative definition of the PoI: any act that might convey to a reasonable actor that he is not presumed innocent of a punishable offence constitutes a PoI interference. The term *act* signifies that this PoI is about behaviour rather than about psychological states.² The phrase 'might convey to a reasonable actor' signifies that the perception³ of the PoI's subject is more important than the intention of the PoI's addressee.⁴ The term *punishable offence* is meant to be an autonomous notion, like *criminal charge* in the ECHR. The PoI still applies when an offence is labelled as administrative law rather than criminal law.

This PoI is so broad that even the usual limiting condition – unless proved guilty according to law – is excluded. Presuming guilt after legal proof is classified as a PoI interference as well. The omnipresence of this limiting condition shows that it is generally accepted to convey guilt after legal proof, but this is still something that should be up for discussion in academia. Excluding all potentially justificatory factors from the definition fosters discussion. Interfering with this PoI is regarded as a *prima facie* wrong. All interferences stand in need of justification and should not be taken for granted from the start. Compare this to the offence definition of homicide. There is a broad and clear rule: do not kill. Well-established exceptions to this rule – justifications and excuses – are not included in the definition. Interfering with the PoI or the homicide prohibition simply means there is a case to answer.⁵

If all themes under discussion are classified as interferences, the discussion need only be about the question what interferences are *justifiable* or *unjustifiable*. This question must be answered in light of Duff's contention that a PoI is an expression of deeper values. But it must also be answered in light of the fact that other values might justify a PoI interference. I will refer to values or interests that might justify a PoI as *pro-PoI values*. Values or interests that might justify a PoI interference will be referred to as *anti-PoI values*. The term *value* is used very broadly: anything that could justify the PoI or a PoI interference is classified as a value. One and the same value can be both pro-PoI and anti-PoI. The value of crime prevention can be pro-PoI: treat ex-offenders as innocents to encourage

2 See Antony Duff, 'Who Must Presume Whom to Be Innocent of What?', this issue, 172-73.

3 If the actor is unaware of the act, knowledge is ascribed. Thus, wire-tapping is a PoI interference.

4 There may be a connection between (reasonable) perception and (perceived) intention.

5 See also Art. 8(2) ECHR: no interference unless necessary ...

them to make a respectable living; it can also be anti-PoI: keep close tabs on them to prevent reoffending.

I believe this approach is better suited to structure the debate than Duff's approach. Duff's approach has three tenets. First, there may be many PoI. Second, a PoI can be defeated (it no longer applies), unqualified (civic PoI) or qualified (roles like *suspect* make the PoI more permissive). Third, acts may be consistent, justifiably inconsistent and unjustifiably inconsistent with a PoI. My first concern is that the system is overly complex. Multiplying the possible categories with however many PoI there may be leads to an expansive PoI universe. My second concern is that this classificatory system is liable to beg the question. Some factors justifying not presuming innocence are positioned outside the PoI: pre-trial detention of serial rapists is inconsistent with the PoI, but justifiable for preventive reasons. Other justificatory factors are incorporated into the PoI: imposing bail on defendants in general, pre-trial detention of defendants who have fled in the past and certain legal presumptions may all be consistent with the PoI.⁶ The act of inclusion or exclusion is guided by normative considerations – some of which may be quite idiosyncratic. The reader is confronted with a predetermined answer to the question whether there is a case to answer.

Duff's multi-PoI approach may be partially prompted by the concern that the issues he wants to discuss are deemed to be outside the scope of *the* PoI. Introducing multiple PoI does avert this problem, but creates others. With multiple PoI, there are just as many scopes to determine. The broad PoI approach grants the same licence to discuss without scope-determining problems. It is also compatible with Duff's contention that roles are crucial for analysing the PoI. Roles are not employed to change a PoI through qualification, but are simply regarded as pro- or anti-PoI values. The broad definition ensures that black-letter lawyers cannot dodge the debate by simply stating that something is consistent with *the* PoI of a legal order.⁷ All assertions must be based on values. Normative legal scholars can theorize about the PoI without worrying about difficult classification issues. The only relevant question is whether a PoI interference is justifiable or unjustifiable.

3 PoI Values and Retributivism

I begin with converting Duff's views into my analytic framework. The following five values function in his analysis. The first value is *desert*. The broad PoI may (or must) be interfered with when it is proved that someone deserves punishment. Desert is an anti-PoI value, but also a pro-PoI value. Once someone has undergone his deserved punishment, his presumed innocence must be restored. The second value, *civic trust*, kicks in when desert is not yet or not anymore an issue.

6 See also s. 5.4.

7 See about the definitional stop H.L.A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford: Clarendon Press, 1968), 4-6.

Alwin A. van Dijk

The trust we owe our fellow citizens is a continuous pro-PoI value. The third value Duff accepts, albeit grudgingly, is *crime prevention*. This value justifies pre-trial detention of serial rapists. The fourth value is the *duty* flowing from a role. The duty of reassurance (posting bail as a defendant) is an anti-PoI value. This is an interesting perspective. Such PoI interferences are usually justified on consequentialist grounds (flight risk). The fifth value, *indiscriminateness*, is also an anti-PoI value. Indiscriminate PoI interferences are deemed less objectionable.

The first-mentioned value of desert only plays a subordinate role in the feature article, but plays a crucial part in Duff's oeuvre. Duff is an influential modern-day retributivist. In his view, punishment must be justified as a mode of moral communication with offenders: it is a species of secular penance.⁸ Punishment seeks to persuade offenders to repent their crimes, to reform themselves and to reconcile themselves with those they have wronged. The justifying aim of secular penance looks suspiciously consequentialist. According to Duff, however, there is no *contingent* relation between punishment and penance. There is an *intrinsic* relationship: it is inherently right to seek to persuade offenders to repent their crimes. Punishment aims to communicate to offenders why their behaviour was wrong. They are not manipulated, but respected as responsible moral agents. In the end, they decide whether they will be persuaded or not. Punishment must even be imposed when persuasion is deemed impossible. This is owed to society (the crime is taken seriously) and to offenders (they deserve the chance to repent).

Retributivist punishment must be proportionate to the offence's seriousness. Proportionality can be violated in two directions: punishment can be too severe or too lenient. In Duff's view, punishment should communicate the correct amount of censure. The communicative purpose bears on the punishment mode. Probation, community service and mediation are well-suited for persuading offenders to repent. Fines are less suitable for penal communication. Imprisonment must be reserved for very serious crimes.

Utilitarianism opposes retributivism because deserved punishment might be dis-utile: should we cause pain without gain? Retributivism opposes utilitarianism because utile punishment might be undeserved: should we punish the innocent for a utility increment? Negative retributivism attempts to counter both objections. Within the bounds of desert, punishment severity should be determined by utilitarian (and retributivist) considerations. *Negative retributivism* entails that the guilty *may* be punished to the extent of their deserts. *Positive retributivism* demands that the guilty *must* be punished to the extent of their deserts. Positive retributivists – like Duff – argue that negative retributivism still denies the guilty the respect due to them as responsible agents. It uses the wrong crime-preventive method: coercion instead of persuasion. Thus, the Hegelian objection that people are treated like dogs remains valid.

8 R.A. Duff, *Punishment, Communication, and Community* (Oxford: Oxford University Press, 2001).

4 Fallibility and the PoI

I will now explore the connection between desert and the PoI's rationale. The question 'when does retributivism *not* need a PoI?' provides a fruitful starting point. This question can be answered by imagining an infallible penal institution, in which all relevant information is perfectly knowable. Imagine we had some kind of non-intrusive desert analyser. This desertometer provides both the amount of desert and the appropriate punishment. Thus, we could always know what constitutes deserved punishment (assuming no PoI prohibits its use).

The infallible retributivist system does not have much use for a PoI because there is no need to *presume* innocence or guilt. This thought experiment shows that the PoI's rationale is closely connected to epistemological considerations. The omniscient state's only remaining task is executing just punishment.⁹ After punishment the desertometer shows that innocence is restored. The roles of suspect, prosecutor, judge and jury cease to exist. Legal concepts like the probable-cause standard, *in-dubio-pro-reo* principle and *nemo-tenetur* principle become unnecessary. Pre-trial detention, bail and trials also become otiose.

In the infallible retributivist system, innocents need never be treated as guilty. This is not true for the infallible utilitarian system. It is improbable that a hedonimeter – a device for measuring and predicting utility – would never indicate it to be utile to punish innocents.¹⁰ Punishment is only contingent on whether it maximizes happiness: guilt and innocence as such are irrelevant. Utilitarianism will punish innocents (if utile); retributivism will never punish innocents. Thus, in an infallible system, innocents are fully protected by retributivism.

Earlier I have distinguished between *justifiable* and *unjustifiable* PoI interferences. Imposing bail on defendants, wiretapping after suspicion or punishing after legal proof may all be justifiable. In a fallible system, the subject of such interferences may in fact be innocent. I refer to such instances as *wrongful* PoI interferences. *Rightful* PoI interferences, in contrast, take place if someone is in fact guilty. Rightful PoI interferences can be unjustifiable and wrongful PoI interferences can be justifiable. The figure depicts the four possible combinations for the PoI interference of punishment.

My use of the terms *rightful* and *wrongful* presumes a hypothetical omniscient perspective. In our fallible system, we can never be absolutely certain whether a

9 Ironically, effective execution of just deserts might lead to massive crime prevention for deterrence-related reasons. People are punished for the 'right' reasons, but refrain from crime for the 'wrong' reasons.

10 Especially, if the state can pretend to only punish the guilty. Availability of desertometers would make this difficult. The term *hedonimeter* is due to Edgeworth. See F.Y. Edgeworth, *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences* (New York: Kelley, 1967), 101.

Alwin A. van Dijk

Figure 1. Combinations of PoI Interferences

	Rightful	Wrongful
Justifiable	Punishing guilty person after proof	Punishing innocent person after proof
Unjustifiable	Punishing guilty person without proof	Punishing innocent person without proof

PoI interference has been rightful or wrongful.¹¹ This is especially true with regard to desert, which poses both conceptual problems (what factors determine desert?) and epistemological problems (how to acquire knowledge about these factors?). All retributivism can do is strive for rightful interferences. Wrongful PoI interferences inevitably occur. Innocents are wiretapped, accused or punished, and guilty persons receive more punishment than they deserve (innocence restored too late). All evils that might legitimately happen in an infallible utilitarian system do actually happen in a fallible retributivist system. This is a compelling reason for always requiring an explicit justification for not presuming innocence.

We live in a fallible system in which many conflicting values compete with each other. Consequentialist values, like utility and prevention, are often portrayed as anti-PoI values. Retributivism, in contrast, is portrayed as providing strong pro-PoI values. Since these retributivist values are regarded as categorical requirements of justice, it is not even deemed legitimate to balance them against consequentialist anti-PoI values.¹² Thus, the impression is created that innocents in a fallible retributivist system are almost as well protected as in an infallible system.

I now turn from theory to practice. The following three sections deal with actual PoI interferences in traffic law: presumptions of guilt, vicarious liability of car owners and coercing non-suspects into proving their sobriety. Retributivists tend to castigate such procedures. I argue, however, that they might be justified on retributivist grounds. In a fallible world, the PoI is not even safe in the hands of retributivists.

5 Rebuttable Presumptions of Guilt

5.1 Prelude: Punishment of Innocents and Evidentiary Standards

Surely, the worst possible wrongful PoI interference is the actual punishment of complete innocents. Unless a retributivist-minded state would refuse to punish anyone, it will inevitably punish innocents. This raises the question whether these *wrongful* PoI interferences could ever be *justifiable*. Thus, retributivism must face the *tu-quoque* argument.¹³ ‘You criticize utilitarianism for punishing inno-

11 Evidence of a wrongful conviction may lead to review. This involves a change in fallible knowledge.

12 See, e.g., Andrew Stumer, *The Presumption of Innocence: Evidential and Human Rights Perspectives* (Oxford: Hart, 2010), 27-51.

13 See David Dolinko, ‘Three Mistakes of Retributivism,’ *UCLA Law Review* 39 (1991-1992): 1623-57.

cents, but are you not guilty of the same charge? Do you not use innocents to achieve the greater good of punishing the guilty?' The retributivist answer to this familiar objection invokes the Doctrine of Double Effect (DDE). 'We do indeed *knowingly* punish innocents, but never *deliberately*!'¹⁴

The DDE has three relevant conditions.¹⁵

1. The act itself must be morally permissible.
2. The good effect is intended; the evil effect is unintended.
3. The good effect must outweigh the bad effect.

I believe with Duff that a clear distinction can be made between intention (direct intent) and foresight (oblique intent). The golden standard is Duff's *test of failure*: 'If my action does not produce an expected effect, will it have been a failure?'¹⁶ Failure indicates intention. Retributivists could argue that punishment of a guilty person is never regarded as a failure. Thus, intention is only directed at punishing the guilty. Is this a valid response to the *tu-quoque* charge? It is not if one believes that directly intending bad consequences is not (much) morally worse than obliquely intending them (*ceteris paribus*).¹⁷ If one does believe there is a large moral difference, the absence of direct intent constitutes the beginning of an answer. However, the DDE's third condition must still be fulfilled: the good effect (punishing the guilty) must outweigh the bad effect (punishing innocents). Even retributivists must revert to consequentialist reasoning.

It would appear that the most important feature of the PoI of legal systems – proof of guilt beyond reasonable doubt or the *in-dubio-pro-reo* principle – is in safe hands with retributivism. Duff does indeed argue for this standard.¹⁸ But this is by no means self-evident. Retributivism has two incompatible duties: the duty to punish the guilty and the duty not to punish innocents. Retributivism does not specify, however, which duty has priority.¹⁹

There are three ways to resolve this conflict. (1) If the duty not to punish innocents is deemed to outweigh the duty to punish the guilty, one might opt for the *in-dubio-pro-reo* principle. This accords with Blackstone's famous maxim that it is better that ten guilty persons escape than that one innocent suffer. (2) If both

14 See R.A. Duff, 'Retributive Punishment – Ideals and Actualities,' *Israel Law Review* 25 (1991): 435-41.

15 See my *Strafrechtelijke aansprakelijkheid heroverwogen: over opzet, schuld, schuldsluitingsgronden en straf* (Apeldoorn: Maklu, 2008), 353. The omitted, subjectively interpreted means-end condition is absorbed by the evil being *unintended*. See Jonathan Bennett, *The Act Itself* (Oxford: Clarendon Press, 1995), 196-200.

16 R.A. Duff, *Intention, Agency and Criminal Liability: Philosophy of Action and the Criminal Law* (Oxford: Basil Blackwell, 1990), 61.

17 See my *Aansprakelijkheid heroverwogen*, 359-65.

18 R.A. Duff, 'Strict Liability, Legal Presumptions, and the Presumption of Innocence,' in *Appraising Strict Liability*, ed. A.P. Simester (Oxford: Oxford University Press, 2005), 133-34.

19 See Russell L. Christopher, 'Deterring Retributivism: The Injustice of "Just" Punishment,' *Northwestern University Law Review* 96 (2001-2002): 910-15. See my *Aansprakelijkheid heroverwogen*, 131-37 for additional references.

Alwin A. van Dijk

duties are deemed equally important, one might opt for a balance-of-probabilities standard. (3) If the duty to punish the guilty is deemed to outweigh the duty not to punish innocents, one might opt for a presumption-of-guilt standard. Guilt is presumed unless innocence is established.

Both the balance-of-probabilities standard and the presumption-of-guilt standard may be compatible with the DDE.²⁰ The intended good effect of the low evidentiary standard is punishment of the guilty. The bad side-effect – punishment of numerous innocents – is foreseen but unintended. Retributivists could reject these standards by arguing that the duty not to punish innocents prevails. They may well meet this persuasive burden. This does not, however, alter the fact that the duty to punish the guilty is an anti-PoI value. This anti-PoI value is only recognized by positive retributivism. The compromise theory of negative retributivism recognizes just one categorical duty: do not punish innocents. This pro-PoI value must be balanced against the anti-PoI value of crime prevention. High evidentiary standards may thwart the law's effectiveness. Still, it is quite possible that negative retributivism would favour a higher evidentiary standard than positive retributivism. The anti-PoI duty to punish the guilty is absent in utilitarianism as well. A relatively high evidentiary standard might be based on the agony of innocent punishees and societal outrage about punishment of innocents.

5.2 Law

(1) *Administrative law.* Minor traffic violations in the Netherlands are usually dealt with by administrative law. When it is established that someone has committed a violation, a fixed administrative sanction is imposed. The accused may lodge an appeal with the public prosecutor. Subsequently, appeal may be lodged with two factual courts. This requires a security deposit. The prosecutor and courts can *moderate* fines based on offence-related circumstances or personal circumstances. They can *annul* a fine when the person did not commit the offence or when offence-related circumstances make it unreasonable to impose a sanction. Although criminal law defences are not as such applicable, justification defences (speeding to bring someone to the hospital) or excuse defences (running a traffic light during a seizure) may nonetheless lead to mitigation or annulment of sanctions.

(2) *Criminal law.* Since 2008, public prosecutors (and others) are authorized to issue *punishment orders* for offences carrying a six-year maximum prison term. This criminal procedure aims to deal with simple cases. The punishment order entails both an act of prosecution and of punishment. Punishment orders may only be issued when guilt has been established. Permissible sanctions are: a fine, a six-month driving disqualification and 180 hours of community service. However, imprisonment is not permitted. The accused can object to a punishment order. If he objects, the regular criminal court procedure applies (including appeal and cassation). The fact that prosecutors establish guilt and impose punishment may be

20 See Christopher, 'Deterring Retributivism,' 910-15.

Figure 2. Evidentiary Scenarios

Offence elements (o)	Defences (d)	Verdict	Reasonable doubt about g
no conviction	–	not guilty	yes
conviction	plausible	not guilty	yes
conviction	implausible	guilty	no
conviction	almost plausible	guilty	yes

shocking to common-law lawyers. It should be borne in mind, however, that Dutch prosecutors fulfil a magistrate-like role: they are trusted to act impartially and justly.

Dutch criminal law distinguishes between felonies and misdemeanours. Traffic violations are generally misdemeanours. The offence definition of misdemeanours contains *actus reus*, but no *mens rea*. The basic structure is: ‘whosoever runs a traffic light is punishable ...’ Intent or negligence is not required. The evidentiary standard for offence elements (*o*) is a *conviction* (firm belief) by the judge. Conviction entails absence of reasonable doubt with regard to *o* (*in-dubio-pro-reo* principle).

After proof of the offence elements, it is presumed that the conduct was wrongful and blameworthy. These presumptions can be rebutted by justifications and excuses. All statutory defences and the extralegal defence *absence of all culpability* are applicable. The latter defence is especially important in traffic cases. It incorporates due diligence and temporary incapacitation to drive (stroke; heart attack). After proof of the offence elements, defences are presumed to be absent (not-*d*). Not-*d* is rebutted if *d* is *plausible*.

I submit that substantive guilt (*g*) is composed out of offence elements (*o*) and the absence of defences (not-*d*). Ergo: $g = o + \text{not-}d$. Although *o* and not-*d* are both part of substantive law, a different evidentiary standard applies (*conviction* with regard to *o* and *plausibility* with regard to *d*). Figure 2 contains four relevant scenarios.

There is no finding of guilt in the first two scenarios. In the third scenario, guilt is established beyond reasonable doubt. The fourth scenario entails a problematic guilty verdict, however. Although the plausibility standard is not met, there seems to be room for reasonable doubt about guilt. E.g., 15% subjective probability of *d* seems insufficient to characterize it as *plausible*. However, 15% subjective probability of *d* logically implies 15% subjective probability of not-*g*.²¹ That seems sufficient for *reasonable doubt*. Thus, the plausibility standard is at odds with the *in-dubio-pro-reo* principle. A *formal conviction* (guilty verdict) need not be based on a *psychological conviction* about all relevant aspects of guilt.

21 Assuming 100% subjective probability of *o*. Less certainty about *o* should induce even more doubt about *g*.

Alwin A. van Dijk

Such discrepancies might be interpreted away on a theoretical level. It seems more fruitful, however, to examine actual case law. The plausibility standard is applied to defences in both administrative and criminal law. Formally, the accused bears no legal burden: defences apply when they are *deemed plausible* rather than *made plausible*. In practice, however, the persuasive burden often rests on the accused. It is he who bears the risk of non-persuasion. In practice, courts are rarely persuaded.²²

5.3 Retributivist Arguments for Strict Procedures

The administrative procedure, punishment order and criminal court procedure all lead to punishment of innocents. Innocents may fail to meet the persuasive burden and unjust punishment orders or administrative fines may go uncontested. These procedures are commonly justified according to consequentialist reasons. Their efficiency and high punishment rate serves deterrence. Retributivists have traditionally objected to such strict procedures by criticizing the consequentialist rationale.²³ This response does not, however, prove the retributivist wrongness of strict procedures altogether. Might it not be argued that strict procedures are sometimes best-suited to serve retributivist goals?

First, it must be established that traffic violations constitute retributivist wrongs. Most traffic violations are implicit endangerment offences. Violating a standard without actually causing danger is criminalized. According to Duff, criminalization of some such offences is indeed justifiable.²⁴ The next question is: what evidentiary standard is appropriate for these wrongs? I submit that traffic offences have two important characteristics. First, it is usually very easy to prove *actus reus*. Speeding, running traffic lights or driving without lights are easily established. Second, it is often very difficult to establish positive proof of the requisite culpability. Culpability factors are to a large extent unknowable for outsiders. I give three examples.

‘I exceeded the speed limit, because I had overlooked the speed sign when I was blinded by headlights.’

‘I ran the traffic light, because I suffered an unforeseeable blackout.’

‘My taillight must have just broken down; I had checked them, as always, when I left.’

These defences could establish moral innocence. However, requiring proof beyond reasonable doubt as to culpability would often amount to a *probatio diabolica*. It is exceedingly difficult for fact-finders to exclude reasonable doubt about not-*d*. This problem might be solved by presuming not-*d* after proof of *o* (*actus*

22 High substantive standards also contribute. An oft-used defence is: absence of *all* culpability.

23 E.g., Duff, ‘Strict Liability,’ 128-29.

24 R.A. Duff, ‘Criminalizing Endangerment,’ in *Defining Crimes: Essays on the Special Part of the Criminal Law*, eds. R.A. Duff & Stuart P. Green (Oxford: Oxford University Press, 2005), 60-62.

reus). In practice, the accused must then make plausible that *d* applies. But this would often amount to a *probatio diabolica* as well. How can someone demonstrate that he was blinded, suffered a blackout or had checked his taillights?

How should positive retributivists deal with such epistemological twilight zones? If positive proof of guilt were required, many guilty people might remain unpunished. Punishment would largely depend on cooperation of defendants. Honest repenters would be punished and liars would escape punishment. The duty to punish the guilty would be seriously compromised. If the burden is imposed on the defendant, however, quite a few innocents might end up being punished. Establishing innocence depends on the fortuitous presence of evidence about *d*.

The latter procedure, which resembles Dutch and other procedures, might be justified on retributivist grounds. First, the duty not to punish innocents is taken seriously. Innocents are never deliberately punished. Furthermore, an innocent person is never knowingly punished.²⁵ Plausibility of innocence negates liability. Second, the duty to punish the guilty is taken seriously. Guilty punishees are always punished deliberately. Third, it is plausible that most convicted are actually guilty. Fourth, money saved by these efficient procedures could be utilized in more serious cases, thus resulting in more justice in those cases. The strict procedure is consistent with the DDE if the expected injustice is outweighed by the expected justice. In that case, it strikes a fair balance between the pro-PoI duty not to punish innocents and the anti-PoI duty to punish the guilty.

5.4 Duff's View

Duff has convincingly argued that presumptions of guilt are occasionally consistent with the reasonable-doubt standard.²⁶ A factory owner can easily prove that he took reasonable steps to make the machinery safe. If he cannot prove this, there is no reasonable doubt. Duff maintains, however, that it is unreasonable to lay such burdens on drivers: driving is an ordinary activity and drivers who fully satisfy their responsibility to drive with due care do not thereby equip themselves to prove their innocence. The last point is certainly true in 'diabolical' cases. Unfortunately, the state's epistemological position is even worse. True defences are almost indistinguishable from false defences. Thus, it seems that nearly all 'diabolical' defences should create reasonable doubt.

It might be objected that presuming innocence in 'diabolical' cases is unwarranted from an epistemological viewpoint. Suppose, *arguendo*, that reliable research showed that 99.9% of blackout defences are false.²⁷ If we cannot properly distinguish between true and false defences, there are empirical reasons to assume a 99.9% probability of guilt in each blackout case. Some might argue that guilt is

25 This is obviously false on a collective level. Numerous unidentifiable innocents are known to be punished.

26 Duff, 'Strict Liability,' 137-43.

27 This might become true if such defences would often create reasonable doubt. Ironically, extending civic trust induces untrustworthiness.

Alwin A. van Dijk

therefore always established beyond reasonable doubt. But this view is difficult to reconcile with Duff's conception of civic trust. It would appear that statistical information by itself provides no valid ground for withdrawal of trust. There is no agent-relative reason for distrust. The reasonableness of doubt depends on the question what information may legitimately be used as evidence. Thus, a high probability of guilt does not necessarily exclude *reasonable* doubt.

In Duff's approach, the anti-PoI duty to punish the guilty conflicts with two pro-PoI values: the duty not to punish innocents *and* the value of civic trust. This approach corresponds with a system in which a relatively high percentage of guilty offenders remain unpunished. Innocents are well-off, but the guilty are denied the opportunity to be persuaded to repent. If less people are persuaded of the wrongness of traffic violations, there will be more traffic violations.²⁸ This has two consequences. First, even more guilty people remain unpunished. Second, there will be more victims due to dangerous driving.

5.5 Conclusion

The 'diabolical' nature of traffic offences poses a challenge for consequentialists and non-consequentialists alike. I hope to have demonstrated that retributivists might go either way on this issue. A more nuanced answer to this question requires more work on the anti-PoI duty to punish the guilty and the pro-PoI value of civic trust.

6 Vicarious Liability of Car Owners

6.1 Law

Earlier I have explained that *administrative fines* for traffic violations may be annulled if non-culpability is deemed plausible. This rule has one exception. If the offence has been committed with a registered motor vehicle and the driver's identity is not immediately established, the fine is imposed on the registered owner. Liability is negated if the owner demonstrates that his vehicle was used against his will, was commercially hired-out or had already been sold. However, lending a car to others does not negate liability. The owner is even liable if he conclusively proves that he did not drive. This administrative procedure was introduced to reduce workload and enable quick execution of fines. It is tremendously successful in these respects. Most of the approximately ten million yearly fines are imposed on registered owners. Most violations are automatically detected by speed and red-light cameras. Nearly all fines are collected within a year.

6.2 Retributivist Arguments for Vicarious Liability

Automatic enforcement methods and the simple liability regime obviously fit well within a consequentialist framework. Violations are easily ascertainable, fines are easily collectable and the high punishment probability reduces harm. Speed cam-

28 Also, because there will be less deterrence as an unintended side-effect of communicative punishment. I doubt, however, whether these consequences carry any weight in Duff's view.

eras may cause a 30-40% reduction in serious injury crashes in these areas.²⁹ The question is, however, whether these benefits are gained at the expense of cherished principles of justice. Does fining the owner for an offence of another not amount to punishment of the innocent?

Scholars had argued that fining non-driving owners violated the ECHR's PoI. However, the European Court of Human Rights (ECtHR) ruled in *Falk* that a complaint to that effect was 'manifestly ill founded.'³⁰ The Court held that the owner is obliged to 'assume the responsibility for his or her decision to allow another person to use his or her car'. Thus, an owner who had lent his car to an offender is *guilty* rather than *innocent*. But this is surely a violation of retributivist principles. Since when does lending your car to a friend merit condemnation? The cynic phrase 'no good deed goes unpunished' comes to mind. Duff has argued that such a formal definition of *guilt* violates the substantive PoI.

'What should matter is whether a defendant has been proved to merit the condemnation that conviction conveys; a statute that so defined "guilt" as to capture those who were clearly innocent of anything that could plausibly count as wrongdoing (...) would not satisfy the substantive presumption that a person is innocent until proved guilty of wrongdoing.'³¹

Does punishing without establishing substantive guilt necessarily amount to an unjustifiable PoI interference according to retributivism? If the procedure serves a retributivist aim, it might be justified via the DDE.³² It could be argued that every act of punishment is aimed at punishing the guilty. Thus, punishment of moral innocents is always unintended. Since lending the car to another is not a valid defence, punishment of a particular, identifiable innocent person will (almost) never occur knowingly.³³ Furthermore, it is plausible that most punish-ees are in fact guilty drivers. Compatibility with the DDE depends on two factors. First, punishment of x guilty people must be deemed to outweigh punishment of y innocents. Second, it must be established that the duty to punish the guilty would be seriously compromised if not driving were a valid defence. This may occur if such systems struggle with attempts to evade identification. Numerous speeding violations in Arizona allegedly remain unpunished: it cannot be proved who is underneath an oft-photographed monkey mask.

The DDE's consequentialist stage is only reached if the evil is *unintended*. Duff has provided a counterargument.

29 See SWOV, 'Speed Cameras: How They Work and What Effect They Have,' (2011). www.swov.nl/rapport/Factsheets/UK/FS_Speed_cameras.pdf.

30 *Falk v. the Netherlands*, appl. no. 66273/01, 19 October 2004.

31 Duff, 'Strict Liability,' 134. But see R.A. Duff, 'Presuming Innocence,' in *Principles and Values in Criminal Law and Criminal Justice: Essays in Honour of Andrew Ashworth*, eds. Lucia Zedner & Julian V. Roberts (Oxford: Oxford University Press, 2012), 63 (arguing for a formal PoI). Thus, the statute only violates the culpability principle.

32 See Christopher, 'Deterring Retributivism,' 904-9.

33 All other justifications and excuses still apply.

Alwin A. van Dijk

'If (...) we make liability for an offence so strict that one who conducts a legitimate activity with all reasonable care may still be guilty (...) we do not simply consent to the mistaken punishment of some who are actually innocent: we intend to convict and punish some whom we believe to be properly innocent (...).'34

But Duff's own test of failure gives another result.³⁵ If a punishee is actually guilty, punishment is a success rather than a failure. Thus, innocents are not deliberately punished. A more convincing counterargument might be: is wilful blindness as to circumstances pertinent to moral innocence reconcilable with retributivism? The work is not done by the wrongful intention to punish innocents, but by the legislator's wrongful intention to promote ignorance about innocence.

Another strategy to reconcile this procedure with retributivism is to deny it involves punishment. The argument is that innocents are not punished and presumed guilty; they are merely 'telished' and presumed 'telishable'.³⁶ This strategy obviously fails: the problem does not disappear by employing the 'definitional stop'.³⁷ Deliberate telishment of innocents stands in need of justification as well.

The Netherlands Supreme Court alluded to a more nuanced version of this argument.³⁸ The Court held that the registered owner is not blamed for the conduct. He only assumes the burden to pay the fine on account of the driver who is guilty of the conduct, so as to reclaim that sum subsequently (if desired). The thrust of this ruling is that the driver – not the owner – is presumed guilty. Since it is up to the owner to reclaim the fine from the guilty driver, one could argue that the fine is still aimed at punishing the guilty driver. The legal right to recover the fine might be construed as a negative condition in the definition of punishment. One might even argue that the owner is empowered by the state with the authority to extract punishment from the guilty. The owner takes the place of the dreaded Fine Collection Agency. If he reclaims the fine, the guilty person is punished after all. But what if the fine is willingly or unwillingly not reclaimed? Does that amount to (unintentional) *punishment* of the innocent? Not necessarily. The owner is in a comparable position as other experiencers of 'collateral punishment' (e.g., spouses of people being jailed or fined). They only experience pain as if being punished.

If fining the non-driving owner cannot be regarded as punishment, this procedure is consistent with retributivism after all. Every act is aimed at punishing the guilty. Many guilty people get their just deserts, either because the owner is guilty or because the guilty driver is punished by the owner. Unintentional punishment

34 R.A. Duff, *Trials and Punishments* (Cambridge: Cambridge University Press, 1986), 159-60.

35 See at n. 16.

36 See John Rawls, 'Two Concepts of Rules,' *Philosophical Review* 64 (1955): 11-12.

37 See Hart, *Punishment and Responsibility*, 4-6 (defining vicarious punishment as a sub-standard case of punishment).

38 ECLI:NL:HR:1993:AC4282.

of innocents only occurs if no blameworthy offence has been committed. This is a valid defence, however. If it is deemed plausible that the driver – owner or non-owner – is blameless, the fine is annulled. Another negative side-effect is the unintentional financial loss when fines are not reclaimed. This is only a minor evil, however. First, it rarely happens. Second, it is not as bad as punishment because the censure is not directed at the non-driving owner. All in all, the expected good effect (punishment of many guilty people) seems to outweigh the expected bad effect (punishment of some innocents and some financial loss).

6.3 *Communicative Punishment*

This administrative procedure may be compatible with some retributivist accounts, but it seems inconsistent with Duff's communicative retributivism. First, most violations are detected anonymously by cameras. Thus, there is no communication of censure directly after the violation. Second, the owner is notified of the fine through an automatically generated letter, which is accompanied by a payment slip. It does not explain why the designated violation merits condemnation. Third, communication of censure is also hampered because it only involves an administrative sanction. The message being: apparently this conduct does not warrant criminal condemnation. Fourth, the state shows no effort to track down the real culprit. The message being: getting paid is more important than being paid by the guilty person. Fifth, fines are liable to communicate the wrong message.³⁹ The message being: the fine is some kind of payment, like using a tollway. Detection probability is weighed against benefits like timesaving. Thus, the law is either violated or obeyed for the wrong reasons.

These objections stem from Duff's retributivist account, but his communicative punishment theory is also valuable for non-retributivist accounts. I favour a Rawlsian theory of punishment, which is neither utilitarian nor retributivist.⁴⁰ The general justifying aim of *crime prevention* (not: utility) is constrained by the extent of the defendant's *control* (not: desert). This theory does not share Duff's objections to coercive means of crime prevention. Deterring traffic violations may be legitimate as long as people have a fair opportunity to avert punishment. It is plausible that fines indeed have a substantial general deterrent effect.⁴¹ But would it not be better – and more effective – if people refrained from traffic violations out of respect for the law's underlying values?⁴²

The moral or educational influence of the administrative procedure leaves much to be desired. The bill's explanatory memorandum actually referred to the traffic violations in question as being 'ethically neutral.'⁴³ Traffic fines suffer from a massive public relations problem: they are generally perceived as an easy way to

39 See Duff, *Punishment, Communication, and Community*, 146-48.

40 See my *Aansprakelijkheid heroverwogen*, 125-64.

41 See SWOV, 'Penalties in Traffic,' (2011). www.swov.nl/rapport/Factsheets/UK/FS_Penalties_in_traffic.pdf.

42 See Tom R. Tyler, *Why People Obey the Law* (New Haven: Yale University Press, 1990).

43 *Kamerstukken II 1987/88*, 20 329, no. 3, 21.

Alwin A. van Dijk

generate revenues. Other sanctions, like the Educational Measure Behaviour and Traffic, are better suited to exert educational influence. But if fines are deemed unavoidable, could communication not be improved? Nozick, another proponent of communicative retributivism, compares punishment to a telegram with the message 'this is how wrong what you did was'.⁴⁴ Perhaps the Fine Collection Agency should do more than stuffing a payment slip in the envelope. I envision a missive with the following thrust.

Dear Citizen,

Traffic violations generally cause a large risk increase. Since the risk in a concrete case is almost impossible to determine – for us and for you – we have had to adopt a universal standard. It cannot be excluded, therefore, that your conduct did not cause a heightened risk. Furthermore, the danger of a heightened risk may still be quite small. However, all these heightened risks taken together cause many fatalities and injuries. Unfortunately, we can only prevent these by punishing everyone who violates the standard. This is why you committed a wrong. We respectfully request your future cooperation in reducing traffic accidents.

Sincerely,

The State

7 Coercing Non-suspects into Proving Their Innocence

7.1 Law

In the Netherlands, the use of investigative powers generally requires a *suspicion of a crime*. In some cases, the lower threshold of *indications of a terrorist crime* applies. The most far-reaching powers, however, are found in traffic law. Designated officers may stop vehicles and order drivers to cooperate with a preliminary breath test. The aim of this power is to control observance of drink-driving laws. The power is not curtailed by a suspicion threshold. Intentional non-cooperation is a felony.

The *preliminary* test may lead to a suspicion of drink-driving. Suspects are obliged to cooperate with an *official* alcohol test. Refusal to cooperate with the official test is a distinct felony. It carries the same penalties as the felony of drink-driving: a three-month maximum prison term and five-year maximum driving disqualification. The results of the official test may be used as evidence of drink-driving.

7.2 Accident Prevention as Justifying Aim

This procedure surely interferes with the broad PoI: it might convey to a reasonable actor that he is not presumed innocent. It embodies the quintessence of civic distrust. Proving one's innocence by providing a 'clean' breath sample is the only way to avoid punishment.

44 Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981), 374.

The procedure can be justified on consequentialist grounds. Approximately 20% of Dutch road fatalities are caused by alcohol.⁴⁵ Reducing drink-driving can prevent many serious road accidents. Random alcohol testing has been particularly effective in case of light offenders. The absence of a suspicion requirement and the penalty for non-cooperation contribute to a high subjective probability of punishment among drink-drivers.

The procedure may also be consistent with my Rawlsian theory of punishment, in which the aim of crime prevention is constrained by the extent of the defendant's control.⁴⁶ Citizens have a fair opportunity to avert punishment. If they abstain from drink-driving and comply with alcohol tests, they will not be punished. In view of the importance of accident prevention and the non-invasive nature of the breathalyser procedure, punishment may not be unfair.

7.3 *Retributivist Arguments for Strict Drink-driving Laws*

With regard to retributivism, there are three difficult questions. First, does drink-driving merit criminal condemnation? Second, is forcing cooperation on pain of punishment justifiable? Third, is testing without suspicion justifiable?

Does the implicit endangerment offence of drink-driving merit criminal condemnation?⁴⁷ Retributivists face two difficulties. First, drink-driving is generally more dangerous than sober driving, but the absolute risk is usually minuscule. A blood alcohol content (BAC) of 0.09% results in a fourfold risk increase. Still, the heightened risk of killing another is considerably smaller than 1/25,000,000 per kilometre. Since there is a plausible connection between desert and probability of harm,⁴⁸ the question is how much punishment is proportionate to such a tiny risk. Second, an intoxicated driver need not even cause more risk than typical sober drivers (reasonable drink-driving).⁴⁹ Statistically, *daily drinkers* with a BAC of 0.09% do not cause an above-average risk. The question is whether non-dangerous drink-driving merits condemnation.

The actual Dutch limit – 0.02% for novices and 0.05% for others – is even more difficult to justify. However, Duff has offered some retributivist arguments that are not dependent on actual danger.⁵⁰ First, citizens have a duty to assure others of safe conduct by visibly following safety-protecting rules. Second, a non-dangerous drink-driver denies fellow-citizenship by claiming superiority. These arguments may go some way, but I seriously doubt that the malevolence that is generally bestowed on drink-driving is consistent with retributivism.

45 See SWOV, 'Driving under the Influence of Alcohol,' (2001). www.swov.nl/rapport/Factsheets/UK/FS_Alcohol_UK.pdf.

46 See at n. 40.

47 See my *Aansprakelijkheid heroverwogen*, 140-44 for additional references.

48 See Andrew von Hirsch & Nils Jareborg, 'Gauging Criminal Harm: A Living-Standard Analysis,' *Oxford Journal of Legal Studies* 11 (1991): 30.

49 See Douglas N. Husak, 'Reasonable Risk Creation and Overinclusive Legislation,' *Buffalo Criminal Law Review* 1 (1998): 599-626.

50 Duff, 'Criminalizing Endangerment,' 59-62.

Alwin A. van Dijk

Is forcing cooperation on pain of punishment justifiable? The possibility of driving disqualification for non-cooperation with the official test suggests punishment is still aimed at censuring drink-driving.⁵¹ Sanction-based cooperation may be compatible with the DDE. Cooperators are only punished if the test shows them to be guilty of drink-driving. Punishment for refusers is harder to justify. Most refusers are probably guilty of two wrongs: drink-driving and non-cooperation.⁵² Determining just deserts for presumed drink-driving is difficult due to unknown BAC. Still, punishment for refusers might be justified. Non-cooperating *drink-drivers* receive a more or less proportionate punishment. Non-cooperating *sober drivers* receive a disproportionate punishment: they should only be punished for non-cooperation. However, the disproportionate part is never inflicted deliberately or knowingly. If the expected good effect (punishment of drink-drivers) is deemed to outweigh the expected bad effect (disproportionate punishment of some non-cooperating sober drivers), this procedure is compatible with the DDE.

Are preliminary alcohol tests without suspicion justifiable? The proportion of alcohol offenders during weekend nights declined from 15% (1973) to 2.4% (2010).⁵³ It appears⁵⁴ that most instances of alcohol testing amount to wrongful PoI interferences. Is such a high number of expected *wrongful* PoI interferences *justifiable* via the DDE? The aim of the procedure is to distinguish between innocent and guilty drivers. Testing of innocents never occurs deliberately: if an expected innocent tests positive for alcohol, the test is not a failure.⁵⁵ The foreseeable (and foreseen) testing of innocents may be justified if the retributivist duty to punish the guilty is taken seriously. Numerous guilty drink-drivers would remain unpunished if suspicion were required. It is generally very difficult to distinguish drink-drivers from sober drivers just by observing driving behaviour. The expected good effect (punishment of many more drink-drivers) could be taken to outweigh the expected bad effect (testing many sober drivers).

Retributivists might even be willing to go further than consequentialists. Since consequentialists are not guided by a duty to punish the guilty, an optimal enforcement level is sufficient. An 'optimal' enforcement level that leaves a large percentage of drink-driving incidents unpunished, may breach the anti-PoI duty to punish the guilty. Thus, more random alcohol tests – and presumably, more wrongful PoI interferences – might be argued for on retributivist grounds.

51 The state shows little effort, however, to discover the truth. Forcefully taking blood from suspects for testing would advance truth-finding, but the intrusive nature is morally problematic. A rebuttable presumption of intoxication is another option.

52 The wrongness of non-cooperation is contentious.

53 See SWOV, 'Driving under Influence.'

54 See at n. 11.

55 See at n. 16.

7.4 *Duff's View*

The flagrant civic distrust underlying the Dutch procedure is in tension with Duff's view. There are, however, three Duffian reasons speaking in its favour. First, random alcohol tests are well-suited to communicate the wrongness of drink-driving to society. Second, the procedure is applied indiscriminately: the distrust should not be taken personally. Third, cooperation might be grounded on a civic duty to assure fellow citizens of one's competence.

8 Conclusion

I have introduced a broad PoI to facilitate a scholarly debate about the justifiability of PoI interferences, using pro- and anti-PoI values. PoI interferences are commonly justified on consequentialist grounds. I have argued, however, that PoI interferences in the investigative, substantive and evidentiary sphere might also be justified on retributivist grounds. The retributivist duty to punish the guilty may be the worst enemy of innocents.