

DANIEL KORMAN

THE FAILURE OF TRUST-BASED RETRIBUTIVISM

(Accepted 29 May 2003)

Many recent attempts at justifying the institution of legal punishment have departed from traditional consequentialist accounts, and have taken the form of retributivist theories. Retributivist theories of punishment are “backwards-looking”, in that they attempt to justify punishment on the basis of events that occurred prior to the time of punishment; whereas consequentialist theories of punishment are “forward-looking” in that they try to justify punishment on the basis of events that are expected to occur as a result of the punishment.¹ Retributive theories commonly appeal to the negative desert acquired by the offender as a result of his crime as a means of justifying legal punishment. But there are less prominent varieties of retributivism that appeal to such considerations as the rights that the offender has forfeited as a result of his crime, the unfairness of the offender’s crime, or the offender’s violation of the victim’s trust. The theory of justification based on this last sort of consideration – which one might call “trust-based retributivism” – has received little critical attention, and will be the main focus of this paper.

The interest in defending the moral permissibility of punishment on the basis of trust violations seems to have stemmed from a series of papers on trust by Annette Baier.² The earliest attempt to justify punishment on trust-based grounds was made by David

¹ This way of characterizing the distinction between consequentialist and retributivist theories was suggested by Michael, Mark A. in “Utilitarianism and Retributivism: What’s the Difference?”, *American Philosophical Quarterly* 29/2 (1992), pp. 173–182.

² Both trust-based retributivist authors I will be considering (Hoekema and Dimock) cite Baier’s “Trust and Anti-Trust”, *Ethics* 96/2 (1986), pp. 231–260. Also, see Baier’s “What do Women Want in a Moral Theory?”, *Nous* 19/1 (1985), pp. 53–63.



Hoekema,³ who claims that trust violations should be punishable just in case the victim enters into the trust relationship involuntarily. I briefly discuss some objections to Hoekema's theory in Section I. The main focus of this paper is Susan Dimock's more thorough defense of trust-based retributivism.⁴ Roughly, Dimock's claim is that the government is justified in punishing criminals insofar as punishment restores the objective reasons that we, as members of a society, have for trusting one another. I argue that Dimock fails, on two counts, to establish that trust-based considerations can justify punishment. First, the same considerations justify the punishment of the legally innocent, and fail to justify the punishment of some offenders. Second, depending on whether one takes a strong or weak reading of the intentional harm involved in punishment, Dimock's theory either (a) fails to explain why it is morally permissible for the government to intentionally harm offenders, or (b) does not really justify punishment at all, but only some alternative to punishment.

Let us begin with a brief, general discussion of what punishment is, and why it stands in need of justification. Dimock offers the following definition of 'punishment', which I will accept for the purposes of this paper with some minor qualifications: "punishment is an avoidable loss intentionally imposed by a legal authority upon an offender for an offense", where an offense is "a violation of a legal rule or principle".⁵ More specifically, this should be taken as a definition of *legal* punishment, since a parent can legitimately be said to have punished his child, even though the parent is not a legal authority. Furthermore, if we wish to understand this as a definition of legal punishment, it is important to understand this definition as implying that the legal authority imposes the loss *in his capacity as a legal authority*; otherwise, the definition construes a legal authority's spanking of his child as an instance of legal punishment.

There are at least two important desiderata for any justification of legal punishment. First, it must be shown that the proposed justification of legal punishment "draws the line" correctly: that is, it justifies

³ Hoekema, David A. "Trust and Obey: Toward a New Theory of Punishment", *Israel Law Review* 25/3 (1991), pp. 332–350.

⁴ Dimock, Susan "Retributivism and Trust", *Law and Philosophy* 16/1 (1997), pp. 37–62.

⁵ Dimock (1997, p. 38).

the punishment of legal offenders, and does not justify punishing (or, intentionally harming) the legally innocent. Second, showing that punishment is morally permissible requires showing that the intentional harm involved in punishment is compatible with, or else can plausibly be taken as an exception to, the compelling moral principle that intentional harm is morally impermissible. The main reason that punishment stands in need of justification is that the harm involved in punishment is *intentional*, as opposed to merely foreseen. ‘Intentional’, in this context, does not simply mean ‘deliberate’. The government engages in deliberate harms to its citizens, such as taxation and quarantine, that are taken to be morally unproblematic. In these cases, the harm is merely *foreseen*, and is not intentional in the relevant sense. (I will return to some of the subtleties involved in distinguishing intended from merely foreseen harms in Section III.) I will show that neither Hoekema’s nor Dimock’s trust-based justifications meet either of these desiderata.

I. DISTINGUISHING PUNISHABLE FROM NON-PUNISHABLE TRUST VIOLATIONS: HOEKEMA

Trust-based retributivism is the theory that the legal system is justified in punishing all and only trust violators (of the right sort). One of the main problems facing a trust-based retributivist theory is that of distinguishing between those violations of trust that ought to be punished, and those for which legal action is not appropriate.⁶ For instance, any reasonable theory of punishment ought to entail that actions such as theft and kidnapping are legally punishable, but that white lies and infidelity are not. But since all of these involve breaches of trust, in order to avoid the absurd conclusion that infidelity ought to be a punishable offense, the trust-based retributivist must supplement his theory with some means by which

⁶ Another potential problem, which I will not pursue here, is that the trust-based retributivist must show that *all* punishable offenses involve trust violations. For instance, consider a devious individual who no one trusts at all. When he commits a crime, no one’s trust has been violated, because no one trusted him in the first place. Nevertheless, his crime may still seem to us to be a punishable offense, and the trust-based retributivist is unable to account for this. Thanks to Matt Tedesco for bringing this to my attention.

to exclude those trust violations that (intuitively) do not warrant legal action. In short, unless the trust-based retributivist can draw the necessary distinction, he will fall victim to the problems of punishing the legally innocent and not punishing the guilty.

One criterion to which one might appeal as a means of distinguishing between punishable and non-punishable breaches of trust is the severity of the harm involved in the trust violation. But both Hoekema and Dimock reject this criterion, and opt for alternative criteria. Hoekema attempts to draw the distinction by appealing to the *voluntariness* with which the victim of a trust violation had entered the trust relationship: “The reason that punishment is inappropriate [in certain cases of trust violation] has to do not with the gravity of the harm caused but with the voluntary character of the trust relation.”⁷ So, for Hoekema, if the individual whose trust has been violated has extended his trust voluntarily, then the trust violation should not be illegal and is not punishable; but when the violation of trust occurs in a situation in which the truster has no choice but to trust the entrusted, then and only then is the breach of trust punishable.

Hoekema argues that taking voluntariness as his criterion allows him to explain why the pharmacist who sells poisonous drugs is fit to be punished, while the grocer who sells spoiled food is not.⁸ One need not trust the grocer not to sell one poisonous food, since one can check for oneself whether some food has gone bad (e.g., with one’s nose and eyes); and if one does trust the grocer, the extension of trust is wholly voluntary, and violations of this trust are not punishable. But since one cannot check for oneself whether the pharmacist’s drugs are lethal, one has no choice but to trust the pharmacist, and therefore the trust relationship is involuntary and violations thereof are punishable. Hoekema argues that this distinction also enables the trust-based retributivist to explain and justify the severity with which the legal system deals with violations of children’s trust, since children have no choice but to trust their parents not to abuse, neglect, or abandon them.⁹

⁷ Hoekema (1991, p. 347).

⁸ Hoekema (1991, p. 346).

⁹ Hoekema (1991, p. 347).

But it can be shown that the involuntariness of one's extension of trust is neither necessary nor sufficient for a trust violation's being punishable. First, consider the case in which a skilled chemist purchases some acne medication from her pharmacist, which the pharmacist has laced with arsenic. The chemist not only knows how to test whether a substance contains arsenic, but also has the means at her disposal. By Hoekema's lights, the chemist's trust relationship with the pharmacist is wholly voluntary (since it is *not* the case that she has no choice but to trust him), and therefore the pharmacist's violation of her trust is not punishable. But, intuitively, what the pharmacist has done should be a punishable offense. So, the involuntariness of a trust relationship is not necessary for a trust violation to be punishable.

Next, consider a case in which a mother promises her child that when she returns from her errands she will have a new toy for him. The child trusts his mother to bring him the toy. By Hoekema's lights, the child has not entered into this trust relationship voluntarily since, confined to his crib, he would be unable to obtain the toy on his own; and therefore, when the mother returns without the toy, she has committed a punishable offense. But, intuitively, breaking a promise to one's child (at least a promise of this sort) should be neither illegal nor punishable. So the involuntariness of the trust relationship is not sufficient for a trust violation to be punishable either.

One might object that the child *did* enter into this trust relationship voluntarily, since he could have refrained from trusting his mother to keep her promise. But while it is true that the child's trust is voluntary in some sense, this cannot be the sense of 'voluntary' which Hoekema has in mind. Otherwise, since the child could refrain from trusting his parents not to abuse and neglect him, and since the chemist could refrain from trusting her pharmacist by not ordering acne medication, child abuse and pill-poisoning would not be punishable on Hoekema's view. So, Hoekema must have some notion of voluntariness on which one's trust can be involuntary even if one can, in principle, refrain from entering the trust relationship. And since Hoekema is unwilling to appeal to the severity of the harm involved in the trust violation (and, I think, rightly so), it is not clear how his theory can avoid the counterexamples given above.

One also might object that a theory of punishment need not justify the punishment of *all* legal offenders, since there are sometimes mitigating circumstances that intuitively excuse an offender from being punished. I wholly agree, and I think that any adequate theory of punishment ought to be able to accommodate mitigating circumstances. The main point behind “drawing the line”-type objections is that theories of punishment ought to match our intuitions about which offenses warrant legal action and which do not. Since in some cases we have intuitions that some offenses are excusable, we should expect a theory of punishment to entail that those offenses should not be punished; but since we have clear intuitions that the pharmacist’s actions do warrant legal action, Hoekema’s theory fails since it entails that this particular offense is *not* punishable.

In addition to justifying the punishment of the legally innocent and failing to justify the punishment of the guilty, Hoekema’s theory is generally unable to justify the intentional harm involved in punishment. Since his theory fails in this respect for roughly the same reason as Dimock’s version of trust-based retributivism, I will postpone discussion of this objection as it affects both theories until Section III. His account may also fall victim to other objections. For instance, Ruth Gavison rejects Hoekema’s theory on the grounds that, by locating the punishable crime in the offenders’ trust violations, we would be punishing him for the wrong reasons: “when I am brutally raped by a total stranger, what is terrible about the offence, and what calls for punishment, is not the violation of my trust. It is the violation of my person, my dignity, and my violation of security.”¹⁰ One may or may not be moved by Gavison’s objection – I, for one, think that it can be overcome by the trust-based retributivist – but the consistency of Hoekema’s theory with the punishment of the innocent and the theory’s inability to justify intentionally harming certain offenders are sufficient to show the theory to be inadequate.

¹⁰ Gavison, Ruth “Punishment and Theories of Justice: Comments on Sadurski and Hoekema”, *Israel Law Review* 25/3 (1991), pp. 351–359.

II. DISTINGUISHING PUNISHABLE FROM NON-PUNISHABLE TRUST VIOLATIONS: DIMOCK

Dimock employs a quite different strategy for distinguishing punishable from non-punishable trust violations. Like Hoekema, she does not think that the relevant difference can be located in the harmfulness of the trust violations.¹¹ Dimock argues that the crucial difference between punishable and non-punishable violations of trust is that the former involve the violation of the *objective* grounds of trust (or, *basic* trust), whereas the latter concern only the *subjective* grounds of trust. On her view, “those violations that make mistrust more objectively reasonable than it would otherwise be, especially mistrust of those other than just the violator, are betrayals with which the law ought to be concerned.”¹² Dimock sees the purpose of the law as that of “maintaining the objective grounds of trust,”¹³ i.e., ensuring that it is objectively reasonable for members of the community to trust one another.

To illustrate this distinction, Dimock has us consider the following scenarios:¹⁴

- (a) The child who does not receive some promised gift.
- (b) The wife who commits adultery.
- (c) The stranger who purposely gives someone wrong directions.
- (d) The child who is physically abused or neglected
- (e) The wife who knowingly infects her husband with HIV.
- (f) The stranger who robs someone.

All six scenarios involve some violation of trust, but (a)–(c) are not punishable, and Dimock agrees that they should not be punishable.

But it is not at all clear that merely distinguishing between objective and subjective violations of trust will enable Dimock to fare any better than Hoekema at correctly drawing the line between those violations of trust that intuitively ought to be illegal and punishable, and those that do not warrant legal action. It seems that, *contra* Dimock, all six of the above cases involve making

¹¹ Dimock (1997, pp. 50–51).

¹² Dimock (1997, p. 51).

¹³ Dimock (1997, p. 45).

¹⁴ Dimock (1997, pp. 49–50).

trust less objectively reasonable. When the child's parents break their promise with impunity, it surely makes it less objectively reasonable – for this child and for children everywhere – to trust their parents. And given Dimock's claim that "violations that make mistrust more objectively reasonable than it would otherwise be . . . are betrayals with which the law ought to be concerned", it follows that the parents' violation of the child's trust is sufficient for them to be punished. Similarly, the case of wife who infects her husband with HIV seems no different in terms of objective and subjective loss of trustworthiness than the wife who infects her husband with the flu. The only difference seems to be the severity of the breach of trust and the degree of harm involved; but Dimock opts not to invoke such considerations in distinguishing between punishable and non-punishable trust violations.

Consider a further example from Dimock meant to illustrate the distinction between objectively and subjectively grounded trust: "even if a woman happens to avoid being abused or sexually assaulted by the men with whom she must live . . . she cannot have objectively grounded reasons to trust men if she knows that men can abuse and assault women with impunity."¹⁵ So, if the law were to allow men to get away with assaulting women, then this woman would lack objective reasons to trust men not to abuse her, despite any subjective reasons she may have to trust them. But isn't this equally true of women whose hearts are broken by insensitive men? Since a man can break a woman's heart with impunity, and since men do in fact often get away with this, women lack objectively grounded reasons to trust men not to break their hearts. So, by Dimock's lights, heart-breakers ought to be prosecuted and punished, in order to reestablish basic trust in the community.

Let us consider one final, especially absurd, situation in which Dimock's proposed justification of punishment entails the punishment of the legally innocent. Consider a drug dealer and his customers. Their interactions are founded upon trust. Now suppose that an informant violates the dealer's trust by ratting him out to the cops. That the informant can violate the dealer's trust with impunity surely makes it less objectively reasonable for drug dealers to trust their customers. So Dimock's theory entails the absurd conclusion

¹⁵ Dimock (1997, pp. 51–52).

that the informant ought to be punished, since this is necessary for returning the community to its previous level of objective reasonableness of trusting – in particular, the reasonableness of trusting people not to inform the cops should they choose to sell drugs.

One may object to the above arguments on the grounds that, on Dimock's definition of punishment, non-offenders cannot be punished, and that it is therefore impossible *by definition* for the legally innocent to be punished. This is essentially right, but the force of the argument against Dimock can be salvaged by putting the point slightly differently. Consider a different possible institution, which one might call "trunishment", where trunishment is an avoidable loss intentionally imposed by a legal authority upon a basic-trust violator for a basic-trust violation. Now the objection can be restated as follows: Dimock's proposed justification of punishment (i.e., that the legal system is morally entitled to intentionally harm trust violators) also justifies the imposition of trunishment, which in turn entails the permissibility of imposing intentional harm on the legally innocent. Whether or not one thinks that the innocent can be *punished*, strictly speaking, Dimock's view still entails the moral permissibility of intentionally harming the legally innocent, via trunishment.

Moreover, Dimock (like Hoekema) faces the problem that her theory fails to justify the punishment of some offenders who, intuitively, ought to be punished. For instance, consider a petty thief, the punishment of whom will only make him angrier and much more likely to commit further trust violations upon his release from prison. Furthermore, suppose that general knowledge that he has been punished will have no deterrent value. In such a case (which may be unlikely, but surely is possible), the community may derive subjective trust – both in their fellow citizens and in the legal system's ability to maintain trust conditions – from the knowledge that the thief has been punished. But their trust will not be objectively grounded, since the thief's anger makes it even less reasonable to trust him now. Dimock is faced with a dilemma. Since punishing the thief will, upon his release, cause the objective levels of trust to drop lower than they would have been had he gone unpunished, Dimock must either say that (1) we are not justified in punishing him, or (2) we are not justified in ever releasing him. If she opts for the first horn, then her theory fails to justify the punishment

of a guilty offender, and again draws the line incorrectly between punishable and non-punishable offenses. If she opts for the second horn, then her theory justifies disproportionate amounts of punishment, since it entails that this thief-with-a-grudge must serve a life sentence.

III. THE PROBLEM OF INTENTIONAL HARM

Dimock, like most authors in the literature, incorporates the notion of intentional harm into her definition of punishment. But it is important to note that there is an ambiguity involved in the notion of intentional harm. Consider a case in which the government quarantines Joe because he catches some highly contagious disease. There is a straightforward sense in which the quarantine is an intentional harm, insofar as the quarantine was deliberately imposed on Joe in full knowledge that he would suffer a loss (of freedom and comfort) as a direct result of the quarantine. But this is not the sense of 'intentional harm' that makes punishment morally problematic, just as quarantining Joe itself is not morally problematic. Let us call this sense on which quarantine counts as an intentional harm the "weak reading" of 'intentional harm'.

The quarantine case may be contrasted with a case in which Sally is imprisoned, and Sally's captors would not have imprisoned her if the imprisonment would not have caused her to suffer. In this second case, the harm is intentional in the "strong sense", since the harm plays an integral role in the imprisonment. This kind of intentional harm is different than the harm involved in the previous case, since all that is intended in the case of the quarantine is that Joe's disease not be spread; and if there were a feasible alternative means of accomplishing this without causing Joe any harm or loss, then the government would presumably opt for this alternative means. The harm is just a foreseen and unfortunate by-product of the quarantine. But in the case of Sally's imprisonment, the harm is strongly intentional. Note that in order for a harm to be intentional in this strong sense, a harm need not be intended for its own sake, but may be intended as a means to some end; whereas, in the quarantine case, the harm is intended neither as a means nor as an end.

Punishment is morally problematic because it essentially involves intentional harm in the strong sense. Proponents of punish-

ment take the harm involved in punishment to be an integral and indispensable aspect of the punishment, and would not be interested in any comparable institution that did not harm offenders. Someone who condones the institution of imprisonment but thinks that it would be better if the offenders did not suffer while imprisoned, is surely not condoning *punishment*, but rather some alternative to punishment. So the defender of punishment must show that intentional harm (in the strong sense) is permissible in the case of legal offenders.

Neither Dimock nor Hoekema provides a satisfactory argument for the permissibility of punishment on the strong reading of ‘intentional harm’; and, I will argue, if they only mean to justify intentional harm on the weak reading, then they are not really justifying punishment at all. The general problem for Dimock is that, even if she can establish that punishment does restore the objective reasonableness of trust, and moreover that punishment is *necessary* for the achievement of this end, she still has to explain why it is permissible to punish trust violators in order to attain this end. Likewise, even if Hoekema can establish that considerations of voluntariness can correctly “draw the line” between offenders and non-offenders, he would still have to explain why these trust violations make it permissible for the government to inflict intentional harm on offenders. Since Hoekema has no discussion of this issue, I will focus exclusively on Dimock’s discussion; but it should be clear that the objections to Dimock’s theory carry over to Hoekema’s theory as well.

Dimock provides two reasons why objective trust violations are sufficient for punishment. (1) “The reason is that punishment is necessary to reaffirm for the members of the community the commitment to basic trust.”¹⁶ That is, punishment is the only way to restore the objective reasonableness of trust to the level it was at prior to the offense. (2) “It is necessary that when trust between members of society has been violated, trust in the law as capable of maintaining the conditions of trust be reaffirmed. Punishment serves this purpose.”¹⁷ That is, punishment is necessary for restoring “meta-trust”, i.e., the community’s objective grounds for trusting

¹⁶ Dimock (1997, p. 53).

¹⁷ Dimock (1997, p. 54).

the government's ability to maintain the objective reasonableness of trust within the community. Is Dimock right that punishment is *the only way* to restore and maintain trust and meta-trust? And even if we grant that restoring and maintaining trust and meta-trust – as a matter of necessity – could not be accomplished except by means of punishing trust violators, the question still remains: why should we think that punishment is a *morally permissible* means of attaining this end?

(In what follows, I focus exclusively on the latter question. But one might wonder whether punishment really is necessary for restoring and maintaining trust conditions. Surely punishment is not *logically necessary*, since we can easily imagine alternative means of restoring and maintaining objective conditions of trust that do not require intentionally harming trust violators. Consider, for instance, a system on which there are laws that threaten to punish trust violators (e.g., with prison sentences), but that when a trust violator is apprehended, he lives a fairly comfortable life in exile; and, at the end of his sentence, he is hypnotized and led to believe that he had spent the time behind bars and utterly miserable. The whole community (including potential offenders) falsely believes that trust violators are punished, and it remains a complete secret that no one is ever actually punished. This scenario is obviously far-fetched, and perhaps all Dimock means to claim is that punishment is the only *feasible* way of restoring and maintaining trust conditions; but then we would at least need to see arguments to the effect that alternatives to punishment such as pure restitution or a purely rehabilitative system are either infeasible or ineffective means of attaining this end.)

Dimock makes no explicit argument in support of the move from the necessity of punishment for restoring trust to the moral permissibility of punishment; but I think some implicit arguments can be teased out (although I am wary of attributing these arguments to Dimock). One sort of argument can be gleaned from Dimock's statement that "if the criminal is not punished, this will demonstrate an unwillingness . . . on the part of the community to maintain the conditions of trust."¹⁸ I suppose the implicit argument might go as follows: since (1) the community is willing to maintain the

¹⁸ Dimock (1997, p. 54).

conditions of trust, and since (2) punishment is necessary for maintaining of the conditions of trust, (3) trust violators must therefore be punished. This argument is not valid as it stands; and, in order to make it so, we must replace the first premise with something along the lines of (1'): the community is willing to do whatever is necessary for the maintenance of the conditions of trust. But now the first premise is far more controversial, because anyone moved by the problem of punishment (i.e., the problem of intentionally harming *anyone*) would be unwilling to intentionally harm offenders, despite a general willingness to maintain trust. For similar reasons, if it turned out AIDS could be cured but only by means of Tuskegee-style experiments, I suspect that most people would be unwilling to support the experiments, even though everyone is generally willing for AIDS to be cured. So, even if we grant that punishment is the only way to maintain trust conditions, the mere desirability of maintaining objective trustworthiness is not sufficient to establish that punishment is a morally permissible means to this desired end.

A second possible argument for the move from trust-maintenance to the permissibility of punishment involves Dimock's claim that "the principal function of a legal system is to create and maintain conditions of trust in a community," and the conditional claim that, if that is its function then trust violators must be punished.¹⁹ But why should anyone accept the conditional claim? Even if we agree with Dimock's claim about the function of the legal system, it is not at all clear how to justify the claim that it is morally permissible (let alone *obligatory!*) for the legal system to perform its function. This argument is reminiscent of a scene in the movie *Speed*, in which Dennis Hopper berates Keanu Reeves for defusing his bombs and not allowing them to fulfill their purpose (i.e., of blowing things up). We can agree that the function of bombs is to explode without agreeing that it is therefore morally permissible or obligatory to allow them to explode. Likewise, it does not follow just from the (alleged) fact that the function of the legal system is to sustain the conditions of trust, that it is morally permissible for the legal system to perform that function.

One might be tempted to respond to this objection by pointing to the absurdity of the implication that it might be immoral to institute

¹⁹ Dimock (1997, p. 37).

and enforce laws. But this apparent absurdity vanishes when we see that something very much like the legal system – call it “the shlegal system” – *would be* morally permitted to perform its function, where the principal function of the shlegal system is to maintain conditions of trust in a community *except when* doing so requires engaging in some moral transgression. So, whereas enforcing the law may sometimes be morally problematic, it would be morally permissible (and perhaps obligatory) to enforce the “shlaw”.

Recall that at the beginning of this section, I mentioned that there is both a strong and a weak reading of ‘intentional harm’, and that Dimock’s justification of punishment fails on either reading. We have already examined why her argument fails to justify punishment on the strong reading; and I will now briefly explain why her argument fails to justify punishment on the weaker reading of ‘intentional harm’. The idea is that punishment intuitively *does* (as a matter of conceptual analysis) involve intentional harm in the stronger sense, and that this is why punishment is morally problematic and stands in need of justification. If all Dimock means to justify is the *deliberate* (i.e., non-accidental) imposition of an avoidable loss by a legal authority upon an offender for a violation of a legal rule or principle, then she really isn’t justifying *punishment*, so much as an alternative to punishment. The point is that, on this weaker reading, Dimock would not be justifying an institution that stands in need of moral justification, and neither would she be justifying the institution that other moral philosophers are trying to justify under the name ‘punishment’.

But even if she is merely justifying an alternative to punishment, her argument still faces various problems. As we saw in Section II, her proposed justification justifies not only deliberately harming offenders, but also deliberately harming the legally innocent. Furthermore, if she is only able to justify non-intentional harm and something like the shlegal system, then she would have to provide some argument as to why her position is superior to other alternatives to punishment, such as pure restitution.²⁰

²⁰ Hoekema does address problems with restitution, but only insofar as it is taken to be a justification of punishment. This is different than pure (i.e., non-punitive) restitution, which is an alternative to punishment.

IV. CONCLUSION

Hoekema and Dimock explore an innovative approach to justifying punishment, in terms of trust violations and the duties of the legal system to restore and maintain the conditions of trust within a community. But, as we have seen, the proposed trust-based justification fails on at least two counts. First, it fails to properly draw the line between offenders and non-offenders insofar as it has absurd implications involving the punishment of the legally innocent and the failure to punish the legally guilty. Second, it fails to justify punishment, insofar as it fails to show that it is permissible to intentionally harm offenders. Finally, since it fails to justify punishment (i.e., intentionally harming offenders), perhaps we ought to construe the theory as justifying an alternative to punishment, in which case its merits would have to be evaluated against other alternatives to punishment such as pure restitution.²¹

REFERENCES

- Baier, A., 'Trust and Anti-Trust', *Ethics* 96/2 (1986): 231–260.
Baier, A., 'What do Women Want in a Moral Theory?', *Nous* 19/1 (1985): 53–63.
Dimock, S., 'Retributivism and Trust', *Law and Philosophy* 16/1 (1997): 37–62.
Gavison, R., 'Punishment and Theories of Justice: Comments on Sadurski and Hoekema', *Israel Law Review* 25/3 (1991): 351–359.
Hoekema, D. A., 'Trust and Obey: Toward a New Theory of Punishment', *Israel Law Review* 25/3 (1991): 332–350.
Michael, M. A., 'Utilitarianism and Retributivism: What's the Difference?', *American Philosophical Quarterly* 29/2 (1992): 173–182.

University of Colorado at Boulder
735A Marine St.
Boulder, CO 80302
USA
E-mail: korman@colorado.edu

²¹ Many thanks to Graham Oddie, Matt Tedesco, the Center for Arts and Humanities, and my anonymous referees. I especially would like to thank David Boonin for his fascinating lectures on the ethics of punishment, which have profoundly influenced every page of this article.

